SFB 649 Discussion Paper 2012-016

Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes

Johanna Kappus*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes*

Johanna Kappus
Institut für Mathematik
Humboldt-Universität zu Berlin
kappus@mathematik.hu-berlin.de

February 15, 2012

Abstract

For a Lévy process X having finite variation on compact sets and finite first moments, $\mu(\,\mathrm{d}x) = x\nu(\,\mathrm{d}x)$ is a finite signed measure which completely describes the jump dynamics. We construct kernel estimators for linear functionals of μ and provide rates of convergence under regularity assumptions. Moreover, we consider adaptive estimation via model selection and propose a new strategy for the data driven choice of the smoothing parameter.

Keywords: Statistics of stochastic processes • Low frequency observed Lévy processes • Nonparametric statistics • Adaptive estimation • Model selection with unknown variance

MSC(2010): 60G51 • 60J75 • 62G05 • 62M05 • 62G20

JEL Classification: C14

1 Introduction

Lévy processes are the building blocks for a large number of continuous time stochastic models with jumps which play an important role, for example, in the modeling of financial data. Let us mention exponential Lévy models (see e.g. [4, 5] and [1, 24]), time changed Lévy processes ([22]) or stochastic volatility models ([21]). Estimating the parameters of a Lévy process is thus

*This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk" and by the Deutsche Akademische Austauschdienst

not only of theoretical relevance, but also an important issue for practitioners. The problem of estimating, nonparametrically, the jump density of a Lévy process, has received considerable attention over the past few years

Depending on the nature of the observations, there exist two fundamentally different approaches to this problem:

When disposing of continuous time observations of the process, the jumps are directly feasible, which suggests to use the observed number of jumps as an estimator of the expected number and apply some smoothing procedure. This approach has been investigated in [14]. When placing oneself in a high frequency model, that is, when assuming that the distance Δ between the observation times tends to zero at a high enough rate, one might discretise this procedure. A large increment within a small time interval will be due to a large jump, so one is eventually able to "see" the jumps. For the details, we refer to [13, 12] and to [6] and [8] and the discussion therein.

In the present setup, when working in a low frequency model, that is, when assuming that Δ is fixed, the jumps are no longer directly feasible so the above approach is no longer possible. Instead, one has to take into account the structural properties of Lévy processes and infinitely divisible laws. In this setting, one faces a more complicated statistical inverse problem. For earlier work on this subject, see [19, 15, 16, 7].

This paper is organized as follows: In Section 2 we introduce the statistical model and assumptions. We define kernel estimators for linear functionals of $\mu(\mathrm{d}x) := x\nu(\mathrm{d}x)$ and provide upper bounds on the corresponding risk. This approach covers typical examples such as point estimation or estimation of integrals over compact sets.

Section 3 is devoted to the problem of the adaptive choice of the smoothing parameter. The interesting point about these considerations is that we consider a model selection problem with unknown variance and this issue is not only of interest in the Lévy model (see [7]), but also a topic of ongoing research in the related field of density deconvolution with unknown distribution of the noise. For most recent work on this subject, we refer to [9].

We propose here a new approach towards this problem. With φ denoting the characteristic function, an estimator $\frac{1}{\widetilde{\varphi}_n}$ of $\frac{1}{\varphi}$ has been introduced in [20]. The key of our analysis lies in the fact that we consider a slight modification of this estimator. This will enable us to make the pointwise control on $\left|\frac{1}{\varphi} - \frac{1}{\widetilde{\varphi}_n}\right|$ which has been proved in [20] uniform on the real line. This will be the key result for dealing with the stochastic penalty term in the model selection procedure.

2 Nonparametric estimation of linear functionals in the Lévy model

2.1 Statistical model and assumptions

A Lévy process $X = \{X_t : t \in \mathbb{R}^+\}$ taking values in $\mathbb{R}$ is observed at discrete, equidistant time points $\Delta, \cdots, 2n\Delta$. We assume throughout the rest of this paper, that the distance Δ between the observation times is fixed.

We shall work under the following structural assumptions on the process X under consideration:

2.1 Assumptions.

(A1) X is of pure jump type.

(A2) X has moderate activity of small jumps in the sense that the following holds true for the Lévy measure ν:

$$\int_{\{|x|\leq 1\}} |x|\nu(\,dx) < \infty. \tag{2.1}$$

(A3) X has no drift component.

(A4) For one and hence for any $t > 0$, X_t has a finite second moment. This is equivalent to stating that

$$\int |x|^2\nu(\,dx) < \infty. \tag{2.2}$$

Imposing the assumptions (A1) and (A2) is equivalent to stating that the process has finite variation on compact sets.

It is well known that under (A1)-(A4), the Lévy-Khintchine representation takes the following special form: The characteristic function of X_Δ is given by

$$\varphi_\Delta(u) := \mathbb{E}\left[e^{iuX_\Delta}\right] = e^{\Delta\Psi(u)}, \tag{2.3}$$

with characteristic exponent

$$\Psi(u) = \int \left(e^{iux} - 1\right)\nu(\,\mathrm{d}x) = \int \frac{e^{iux} - 1}{x} x\nu(\,\mathrm{d}x). \tag{2.4}$$

(a proof can be found, for example, in [23]). The process is thus fully described by the signed measure $\mu(\,\mathrm{d}x) := x\nu(\,\mathrm{d}x)$, which is finite thanks to (2.1) and (2.2).

We are interested in the problem of estimating some linear functional of μ. That is, given some function or distribution f, the parameter of interest is

$$\theta := \langle f, \mu\rangle := \int f(x)\mu(\,\mathrm{d}x). \tag{2.5}$$

To simplify the problem and avoid a general discussion about distributions, we assume that one of the following conditions is met:

(F1) f *is a function in* $\mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^\infty(\mathbb{R})$.

(F2) For some $y \in \mathbb{R}\setminus\{0\}$, f *is the Dirac distribution* $f = \delta_y$.

In the latter case, we formulate the following additional assumption on μ which makes the problem well defined:

(A5) For some open interval $D = (d_1, d_2)$ *with* $y \in D$, *the restriction* $\mu\big|_D$ *possesses a continuous Lebesgue density* g_D.

That is, the parameter of interest is the density g of μ, evaluated at y.

2.2 Estimation procedure and risk bounds

In a low frequency model, the jumps of a Lévy process are not directly feasible, so we have to take into account the structural properties of infinitely divisible laws to infer the underlying jump dynamics.

Using formula (2.4), we see that the Fourier transform of μ can be recovered by derivating the characteristic exponent:

$$\Psi'(u) = \frac{\partial}{\partial u}\int (e^{iux} - 1)\nu(\,\mathrm{d}x) = i\int e^{iux} x\nu(\,\mathrm{d}x) = i\mathcal{F}\mu(u). \tag{2.6}$$

We do thus have

$$\mathcal{F}\mu(u) = \frac{\frac{1}{\Delta}\varphi'_\Delta(u)}{i\varphi_\Delta(u)}. \tag{2.7}$$

Under mild regularity assumptions, we can express the parameter of interest in the Fourier domain using the Plancherel formula:

$$\theta = \int f(x)\mu(\,\mathrm{d}x) = \frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\mu(u)\,\mathrm{d}u. \tag{2.8}$$

Together with formula (2.7), this yields

$$\theta = \frac{1}{2\pi}\int \mathcal{F}f(-u)\frac{\frac{1}{\Delta}\varphi(u)}{i\varphi_\Delta(u)}\,\mathrm{d}u. \tag{2.9}$$

These formulae suggest to estimate θ by Fourier methods, replacing the characteristic function as well as its derivative by their empirical counterparts.

Since the increments $Z_{\Delta,j} := X_{j\Delta} - X_{(j-1)\Delta}$, $j = 1, \cdots, 2n$ of X form i.i.d. copies of X_Δ, we can define the empirical versions of φ_Δ and φ'_Δ as follows:

$$\widehat{\varphi}_{\Delta,n}(u) := \frac{1}{n}\sum_{j=1}^{n} e^{iuZ_{\Delta,j}} \tag{2.10}$$

and

$$\widehat{\varphi}'_{\Delta,n}(u) := \frac{1}{n} \sum_{j=n+1}^{2n} iZ_{\Delta,j} e^{iuZ_{\Delta,j}}. \tag{2.11}$$

Moreover, the empirical characteristic function appearing in the denominator is replaced by its truncated version, setting

$$\frac{1}{\widetilde{\varphi}_{\Delta,n}(u)} := \frac{1(\{|\widehat{\varphi}_{\Delta,n}| \geq (\Delta n)^{-1/2}\})}{\widehat{\varphi}_{\Delta,n}(u)}. \tag{2.12}$$

This approach is originally due to Neumann (see [20]).

In case that f has integrable Fourier transform, we are in a position to define a direct plug-in-estimator:

2.2 Definition. Assume that (A1)-(A4) are satisfied and that (F1) is met. Moreover, assume that $\mathcal{F}f \in \mathrm{L}^1(\mathbb{R})$. Then we set

$$\widehat{\theta}_{\Delta,n} := \frac{1}{2\pi} \int \mathcal{F}f(-u) \frac{\frac{1}{\Delta}\widehat{\varphi}'_{\Delta,n}(u)}{i\widetilde{\varphi}_{\Delta,n}(u)} \,\mathrm{d}u. \tag{2.13}$$

The integral appearing in (2.13) is well defined since $|\widehat{\varphi}'_{\Delta,n}|$ as well as $\left|\frac{1}{\widetilde{\varphi}_{\Delta,n}}\right|$ are by definition bounded above and $\mathcal{F}f$ is integrable by assumption.

On the other hand, when being interested in point estimation, $\mathcal{F}f$ is certainly not integrable and the integral appearing in (2.13) generally fails to converge. For this reason, we have to introduce an additional smoothing procedure. This leads to defining kernel estimators:

2.3 Definition. Assume that (A1)-(A4) are satisfied and that $(F1)$ or both, $(F2)$ and $(A5)$ are met. Let a continuous kernel K be given such that for arbitrary $h > 0$, $\mathcal{F}\,\mathrm{K}_h\,\mathcal{F}f(-\bullet)$ is integrable. Then we define for a bandwidth $h > 0$:

$$\widehat{\theta}_{\Delta,h,n} := \frac{1}{2\pi} \int \mathcal{F}f(-u) \frac{\frac{1}{\Delta}\widehat{\varphi}'_{\Delta,n}(u)}{i\widetilde{\varphi}_{\Delta,n}(u)} \mathcal{F}\,\mathrm{K}\,(hu) \,\mathrm{d}u. \tag{2.14}$$

This definition is meaningful since boundedness of $\left|\frac{\widehat{\varphi}'_{\Delta,n}}{\widetilde{\varphi}_{\Delta,n}}\right|$ and integrability of $\mathcal{F}\,\mathrm{K}_h\,\mathcal{F}f(-\bullet)$ guarantee that the integral in (2.14) is well defined and finite.

We can proof the following bound on the risk of $\widehat{\theta}_{\Delta,h,n}$:

2.4 Theorem. *Let the assumptions which are summarized in Definition 2.3 be satisfied. Assume, moreover, that for arbitrary $h > 0$,*

$\mathcal{F}\mathrm{K}_h\frac{\mathcal{F}f}{\varphi_\Delta}\in \mathrm{L}^1(\mathbb{R})\cap\mathrm{L}^2(\mathbb{R})$. *Then we can estimate*

$$\begin{aligned}
&\mathbb{E}\left[\left|\theta-\widehat{\theta}_{\Delta,h,n}\right|^2\right] \qquad (2.15)\\
\leq\ & 2\left|\int f(x)\mu(\,dx)-\int f(x)\,(\mathrm{K}_h*\mu)\,(\,dx)\right|^2\\
+\ & \frac{T^{-1}}{2\pi^2}\left\{C_1\int|\mathcal{F}\mathrm{K}\,(hu)\,|^2\left|\frac{\mathcal{F}f(-u)}{\varphi_\Delta(u)}\right|^2\,du\wedge C_2\left(\int|\mathcal{F}\mathrm{K}\,(hu)\,|\left|\frac{\mathcal{F}f(-u)}{\varphi_\Delta(u)}\right|\,du\right)^2\right\},
\end{aligned}$$

with $2T=2\Delta n$ *denoting the time horizon and with constants*

$$C_1=C\left(\int|\Psi''(x)|\,dx+2\int|\Psi'(x)|^2\,dx\right)\leq\infty \qquad (2.16)$$

and

$$C_2=C\left(\|\Psi''\|_\infty+2\|\Psi'\|_\infty^2\right)<\infty, \qquad (2.17)$$

where C *is some universal positive constant.*

The assumption that $\mathcal{F}\mathrm{K}_{\frac{\mathcal{F}f}{\varphi_\Delta}}$ is integrable and square integrable depends on the unknown characteristic function $\frac{1}{\varphi_\Delta}$. However, we can always ensure that this assumption is met by choosing a kernel function which has compact support in the Fourier domain.

Next, we obtain the following upper bound on the risk of the estimator $\widehat{\theta}_{\Delta,n}$, which is defined without any additional smoothing procedure:

2.5 Theorem. *Assume that* $\mathcal{F}f\in\mathrm{L}^1(\mathbb{R})$. *Let* $\widehat{\theta}_{\Delta,n}$ *be defined as in Definition 2.2. Then we can estimate for* arbitrary $m\geq 0$*:*

$$\begin{aligned}
&\mathbb{E}\left[\left|\theta-\widehat{\theta}_{\Delta,n}\right|^2\right]\\
\leq\ & \frac{1}{2\pi^2}\left\{C_1\int\limits_{\{|u|>\pi m\}}|\mathcal{F}f(-u)|^2\,du\wedge C_2\left(\int\limits_{\{|u|>\pi m\}}|\mathcal{F}f(-u)|\,du\right)^2\right\}\\
+\ & \frac{T^{-1}}{2\pi^2}\left\{C_1\int\limits_{\{|u|\leq\pi m\}}\frac{|\mathcal{F}f(-u)|^2}{|\varphi_\Delta(u)|^2}\,du\wedge C_2\left(\int\limits_{\{|u|\leq\pi m\}}\frac{|\mathcal{F}f(-u)|}{|\varphi_\Delta(u)|}\,du\right)^2\right\},
\end{aligned}$$

with constants C_1 *and* C_2 *defined as in Theorem 2.4.*

It is interesting to note that the estimator $\widehat{\theta}_{\Delta,n}$, which is defined without any additional smoothing procedure can be understood as the constructive analogue of the minimum distance estimator which has been proposed in [19]. The clear advantage is that our estimator can be calculated directly from the data and does not require an abstract minimization procedure over spaces of measures, which is certainly comfortable in applications.

2.3 Rates of convergence

In this section, we investigate the rates of convergence which can be derived form the upper risk bounds given in Theorem 2.4 and Theorem 2.5 under the assumption that the signed measure μ, which describes the jump dynamics of the underlying Lévy process belongs to some prescribed smoothness class.

Let us introduce the following abstract nonparametric classes:

2.6 Definition.

(i) We denote by $\mathcal{F}(\beta, \rho, C_f, C'_f, c_f, c'_f)$ the class of functions f such that for any $u \in \mathbb{R}$:

$$C_f(1+|u|)^{-\beta} \exp\left(-c_f|u|^\rho\right) \leq |\mathcal{F}f(u)| \leq C'_f(1+|u|)^{-\beta} \exp\left(-c'_f|u|^\rho\right).$$

If $\rho = 0$ and $\beta > 0$, the functions in $\mathcal{F}(\beta, \rho, C_f, C'_f, c_f, c'_f)$ are called *ordinary smooth*. For $\rho > 0$, they are called *supersmooth*.

(ii) Given $a > 0$, let $\langle a \rangle := \sup\{k \in \mathbb{N} : k < a\}$. For an open subset $D \subseteq \mathbb{R}$, we denote by $\mathcal{H}_D(\alpha, L, R)$ the class of functions f such that $\sup_{x \in D} |f(x)| \leq R$, $f|_D$ is $\langle a \rangle$ times continuously differentiable and we have

$$\sup_{\substack{x,y \in D \\ x \neq y}} |f^{(\langle a \rangle)}(x) - f^{(\langle a \rangle)}(y)| \leq L|x-y|^{a - \langle a \rangle}.$$

The functions belonging to $\mathcal{H}_D(a, L, R)$ are called locally *Hölder regular* with index a.

(iii) For $a, M \geq 0$, the *Sobolev class* $\mathcal{S}(a, M)$ consists of all square integrable functions, for which

$$\int (1+|u|^2)^a |\mathcal{F}f(-u)|^2 \, \mathrm{d}u \leq M \tag{2.18}$$

holds. For negative indices, we are still in a position to define corresponding Sobolev classes. The objects collected in $\mathcal{S}(a, M)$ for $a < 0$ need no longer be square integrable functions, but are those tempered distributions for which (2.18) holds true.

We start by providing rate results under global regularity assumptions on the test function f and on μ, measured in a Sobolev sense. Let us first recall the following definition:

2.7 Definition. A kernel K is called a *k-th order kernel*, if for all integers $1 \leq m < k$,

$$\int x^m K(x) \, \mathrm{d}x = 0 \tag{2.19}$$

and moreover,

$$\int |x|^k |\, \mathrm{K}(x)| \, \mathrm{d}x < \infty. \tag{2.20}$$

Equation (2.19) is equivalent to stating that the derivatives $(\mathcal{F}\,\mathrm{K})^{(m)}(0)$ vanish for $m = 1, \cdots, \langle k \rangle$.

Let us first have a look at the approximation error which results from smoothing with some kernel function K:

2.8 Lemma. *Assume that for some real valued s and some positive constant M_f, $f \in \mathcal{S}(s, M_f)$. Assume, moreover, that for some $a > -s$, $\mu \in \mathcal{S}(a, M_\mu)$. Let* K *be chosen such that either* K *is the sinc kernel or* K *has order $a + s$ and $\mathcal{F}\,\mathrm{K}$ is Hölder-regular with index $a + s$. Then we can estimate*

$$\left| \int f(x)\mu(\,dx) - \int f(x)(\mathrm{K}_h * \mu)(x)\,dx \right|^2 \leq C_b h^{2a+2s} =: b_h. \tag{2.21}$$

with some C_b depending on M_f, M_μ and on K.

Next, we have the following bound on the error in the model:

2.9 Lemma. *Assume that $\mathcal{F}\,\mathrm{K}$ is supported on $[-\pi, \pi]$. Assume, moreover, that $f \in \mathcal{S}(s, M_f)$ and that for positive constants C_φ and c_φ,*

$$\forall u \in \mathbb{R} : |\varphi_\Delta(u)| \geq (1 + C_\varphi |u|)^{-\Delta\beta} \exp\left(-\Delta c_\varphi |u|^\rho\right). \tag{2.22}$$

Let

$$\sigma_h^2 := \frac{C_1}{2\pi^2} \int |\mathcal{F}\,\mathrm{K}(hu)|^2 \left| \frac{\mathcal{F}f(-u)}{\varphi_\Delta(u)} \right|^2 du \wedge \frac{C_2}{2\pi^2} \left(\int |\mathcal{F}\,\mathrm{K}(hu)| \left| \frac{\mathcal{F}f(-u)}{\varphi_\Delta(u)} \right| du \right)^2$$

Then we have $\sigma_h^2 \leq v_{\Delta,h}$ with

$$\begin{aligned} v_{\Delta,h} \quad := \quad \frac{C_v}{2\pi^2} \Bigg\{ & C_1 \sup_{\{|u| \leq \frac{\pi}{h}\}} (1 + |u|)^{2\Delta\beta - 2s} \exp\left(\Delta c_\varphi |u|^\rho\right) \\ & \wedge\ C_2 \int_{\{|u| \leq \frac{\pi}{h}\}} (1 + |u|)^{2\Delta\beta - 2s} \exp\left(2\Delta c_\varphi |u|^\rho\right) du \Bigg\} \end{aligned}$$

and some constant C_v depending on C_φ and M_f.

Now, let us introduce the following abstract nonparametric classes of signed measures:

2.10 Definition. Let $\mathcal{M} := \mathcal{M}(\bar{C}_1, \bar{C}_2, C_\varphi, c_\varphi, \beta, \rho, a, M_\mu)$ be the collection of finite signed measures μ, such that the following holds:

(i) There is a Lévy process X, for which (A1)-(A4) are satisfied, such that $\mu(\,\mathrm{d}x) = x\nu(\,\mathrm{d}x)$.

(ii) For the characteristic function

$$\varphi(u) := \exp\left(\int \frac{e^{iux}-1}{x}\mu(\,\mathrm{d}x)\right) \tag{2.23}$$

of X_1, the following holds:

$$\forall u \in \mathbb{R}: \ |\varphi(u)| \geq (1 + C_\varphi |u|)^{-\beta} e^{-c_\varphi |u|^\rho}. \tag{2.24}$$

(iii) For C_1 and C_2 defined as in (2.16) and (2.17), we have $C_1 \leq \bar{C}_1$ and $C_2 \leq \bar{C}_2$.

(iv) μ is contained in the Sobolev class $\mathcal{S}(a, M_\mu)$.

Let $\mathbb{P}_\mu = \mathbb{P}^{X_1}$ be the infinitely divisible law with characteristic function φ defined by 2.23 and $\mathbb{E}_\mu$ the expectation with respect to $\mathbb{P}_\mu$.

We can now provide rates of convergence, uniformly over those nonparametric classes:

2.11 Theorem. *Assume that $f \in \mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^\infty(\mathbb{R})$ and $f \in \mathcal{S}(s, M_f)$. Consider the nonparametric class $\mathcal{M} := \mathcal{M}(\bar{C}_1, \bar{C}_2, C_\varphi, c_\varphi, \beta, \rho, a, M_\mu)$ with $a > -s$. For $h > 0$, let $\widehat{\theta}_{\Delta,h,n}$ be defined by (2.14). Assume that the conditions on the kernel function which are summarized in Lemma 2.8 and Lemma 2.9 are met. Let b_h and $v_{\Delta,h}$ be defined as in Lemma 2.8 and Lemma 2.9. Then, selecting $h^* = h^*_{\Delta,n}$ as the minimizer of $b_h + T^{-1}v_{\Delta,h}$, we find that*

$$\sup_{\mu \in \mathcal{M}} \mathbb{E}_\mu \left[\left|\theta - \widehat{\theta}_{\Delta,h^*,n}\right|^2\right] = O\left(r_{\Delta,n}\right) \tag{2.25}$$

with $(r_{\Delta,n})$ denoting the sequences which are summarized in the following table:

	$\bar{C}_1 < \infty$		$\bar{C}_1 = \infty$	
$\rho = 0$	$s \geq \Delta\beta$	T^{-1}	$s \geq \Delta\beta + \frac{1}{2}$	T^{-1}
	$s < \Delta\beta$	$T^{-\frac{2a+2s}{2a+2\Delta\beta}}$	$s < \Delta\beta + \frac{1}{2}$	$T^{-\frac{2a+2s}{2a+2\Delta\beta+1}}$
$\rho > 0$		$(\frac{\log T}{\Delta})^{-\frac{2a+2s}{\rho}}$		$(\frac{\log T}{\Delta})^{-\frac{2a+2s}{\rho}}$

Let us compare this result to the rates of convergence which can be obtained for the estimator $\widehat{\theta}_{\Delta,n}$, which is defined without an additional smoothing procedure.

2.12 Theorem. *Let $f \in \mathcal{S}(s, M_f)$ for some $s > \frac{1}{2}$. Consider the non-parametric class $\mathcal{M} := \mathcal{M}(\bar{C}_1, \bar{C}_2, C_\varphi, c_\varphi, \beta, \rho, a, M_\mu)$. Let $\widehat{\theta}_{\Delta,n}$ be defined by (2.13). Then we find that*

$$\sup_{\mu \in \mathcal{M}} \mathbb{E}_\mu \left[\left| \theta - \widehat{\theta}_{\Delta,n} \right|^2 \right] = O\left(r_{\Delta,n} \right), \tag{2.26}$$

with $(r_{\Delta,n})$ collected in the following table:

	$\bar{C}_1 < \infty$		$\bar{C}_1 = \infty$	
$\rho = 0$	$s > \Delta\beta$	T^{-1}	$s > \Delta\beta + \frac{1}{2}$	T^{-1}
	$s = \Delta\beta$	T^{-1}	$s = \Delta\beta + \frac{1}{2}$	$(\log T)T^{-1}$
	$s < \Delta\beta$	$T^{-\frac{2s}{2\Delta\beta}}$	$s < \Delta\beta + \frac{1}{2}$	$T^{-\frac{(2s-1)}{2\Delta\beta}}$
$\rho > 0$		$\left(\frac{\log T}{\Delta}\right)^{-\frac{2s}{\rho}}$		$\left(\frac{\log T}{\Delta}\right)^{-\frac{2a+2s}{\rho}}$

Examples

(i) For Compund Poisson processes, the absolute value of the characteristic function φ is bounded below. Consequently, (2.24) is satisfied with $\beta = 0$ and $\rho = 0$. If the test function is contained in the Sobolev class $\mathcal{S}(s, M_f)$ with $s > \frac{1}{2}$, Theorem 2.12 immediately tells us that $\widehat{\theta}_{\Delta,n}$ attains the parametric rate.

(ii) For Gamma processes with parameters β and λ, the characteristic function is given by

$$\varphi(u) = \left(1 - \frac{i}{\lambda} u\right)^{-\beta}.$$

From this we conclude that for test functions $f \in \mathcal{S}(s, M_f)$ with $s > \frac{1}{2}$, the estimator $\widehat{\theta}_{\Delta,n}$ attains the parametric rate, provided that $\Delta < \frac{s - \frac{1}{2}}{\beta}$.

(iii) A tempered α-stable law is constructed by multiplying the Lévy measure of a α-stable law with a decreasing exponential. The activity of small jumps is the same as for α-stable laws, so the process has finite variation on compacts if $\alpha < 1$. The characteristic function decays exponentially, with $\rho = \alpha$, so the rates of convergence are logarithmic. For the exact parameters, we refer to Section 4.5 in [10].

We recover in Theorem 2.12 the rates of convergence which have been derived for the minimum distance estimator. This confirms the analogy between the constructive estimator defined by (2.13) and the estimator proposed in [19].

Theorem 2.11 suggests that better rates of convergence can be obtained, under regularity assumptions on μ, when applying some kernel smoothing procedure. However, we must be careful about the fact that $\mu(\,\mathrm{d}x) = x\nu(\,\mathrm{d}x)$ cannot possess a globally smooth Lebesgue density, unless we are in the Compound Poisson case. In the case of infinite jump activity, we will always have a point of discontinuity at zero. Consequently, when considering test functions with integrable Fourier transform which do not vanish at the origin, the gain in the rate which results from kernel smoothing is small and one might prefer $\widehat{\theta}_{\Delta,n}$ in applications.

The situation changes if f is bounded away from the origin. In this case, one has to localize the procedure, working with some kernel function K which decays fast enough. The appropriate concept to take into account is no longer global Sobolev regularity but *local* regularity round the point or interval of interest, measured in a Hölder sense.

We can give the following bound on the approximation error under local regularity assumptions on μ and f:

2.13 Lemma. *Let f be compactly supported with* $\mathrm{supp}(f) := [a,b] \subseteq \mathbb{R}\setminus\{0\}$ *and assume that for some $s \in \mathbb{N}$,*

$$\forall u \in \mathbb{R}: \ |\mathcal{F}f(u)| \leq C_f(1+|u|)^{-s}.$$

Assume that for some bounded open set $D = (d_1, d_2) \supseteq [a,b]$, $\mu\Big|_D$ possesses a Lebesgue density $g_D \in \mathcal{H}_D(a,R,L)$. Let K *have order $a+s$ and assume that for some positive constant C_K, we have*

$$\forall z \in \mathbb{R}: |\,\mathrm{K}(z)| \leq C_K(1+|z|)^{-a-s-1}. \tag{2.27}$$

Then we can give the following bound on the approximation error:

$$\left|\int f(x)\mu(\,dx) - \int f(x)(\mathrm{K}_h * \mu)(x)\,dx\right|^2 \leq C_b h^{2a+2s} \tag{2.28}$$

with a positive constant C_b depending on K*, a, b, D, R, and L.*

The following result is in analogy with Lemma 2.9. However, we need to pay attention to the fact, that the definition of the smoothness parameter s is now slightly different.

2.14 Lemma. *In the situation of the preceding lemma, assume that for positive constants C_φ and c_φ, we have*

$$\forall u \in \mathbb{R}: \ |\varphi(u)| \geq (1+C_\varphi|u|)^{-\beta}e^{-c_\varphi|u|^\rho}. \tag{2.29}$$

Assume, moreover, that $\mathcal{F}\mathrm{K}$ *is supported on* $[-\pi, \pi]$*. Then, with* σ_h^2 *defined as in Lemma 2.9, we have* $\sigma_h^2 \leq v_{\Delta,h}$ *with*

$$v_{\Delta,h} := \frac{C_v}{2\pi^2}\Bigg\{ C_1 \int\limits_{|u|\leq\frac{\pi}{h}} (1+|u|)^{2\Delta\beta-2s} \exp\left(2\Delta c_\varphi |u|^\rho\right) du \\ \wedge C_2 \left(\int\limits_{|u|\leq\frac{\pi}{h}} (1+|u|)^{\Delta\beta-s} \exp\left(\Delta c_\varphi |u|^\rho\right) du \right)^2 \Bigg\},$$

where C_v *is a positive constant depending on* C_f *and* C_φ*.*

We consider now the following class of locally Hölder regular measures:

2.15 Definition. Let $\mathcal{M} := \mathcal{M}(\bar{C}_1, \bar{C}_2, C_\varphi, c_\varphi, \beta, \rho, a, D, L, R)$ be the collection of finite signed measures μ, such that the following holds: The items (i)-(iii) from Definition 2.10 are true and

(iv) $\mu\Big|_D$ possesses a Lebesgue density $g_D \in \mathcal{H}_D(a, L, R)$.

The rate results which can be derived from Lemma 2.13 and Lemma 2.14 are summarized in the following theorem:

2.16 Theorem. *Let the assumptions of Lemma 2.13 and Lemma 2.14 be satisfied. Consider the nonparametric class* $\mathcal{M} = \mathcal{M}(\bar{C}_1, \bar{C}_2, C_\varphi, c_\varphi, \beta, \rho, a, D, L, R)$ *defined in 2.15. Let* h^* *be selected as the minimizer of* $b_h + T^{-1}v_{\Delta,h}$*. Then we find that*

$$\sup_{\mu\in\mathcal{M}} \mathbb{E}_\mu \left[\left| \theta - \widehat{\theta}_{h^*,n} \right|^2 \right] = O\left(r_{\Delta,n}\right) \tag{2.30}$$

with the rates $r_{\Delta,n}$ *collected in the following table:*

	$\bar{C}_1 < \infty$		$\bar{C}_1 = \infty$	
$\rho = 0$	$s > \Delta\beta + \frac{1}{2}$	T^{-1}	$s > \Delta\beta + 1$	T^{-1}
	$s = \Delta\beta + \frac{1}{2}$	$(\log T)\, T^{-1}$	$s = \Delta\beta + 1$	$(\log T) T^{-1}$
	$s < \Delta\beta + \frac{1}{2}$	$T^{-\frac{2s+2a}{2\Delta\beta+2a+1}}$	$s < \Delta\beta + 1$	$T^{-\frac{2a+2s}{2a+2\Delta\beta+2}}$
$\rho > 0$		$\left(\frac{\log T}{\Delta}\right)^{-\frac{2a+2s}{\rho}}$		$\left(\frac{\log T}{\Delta}\right)^{-\frac{2a+2s}{\rho}}$

Examples

(i) For point estimation, we have $|\mathcal{F}f(u)| = |\mathcal{F}\delta_y(u)| = |e^{iuy}| \equiv 1$, so (2.29) is met with $s = 0$ and $C_f = 1$. Under the above assumptions on the local smoothness and on the kernel function, we end up with the polynomial rate $T^{-\frac{2a}{2\Delta\beta+2a+1}}$ in case that φ decays polynomially and with the logarithmic rate $\left(\frac{\log T}{\Delta}\right)^{-\frac{2a}{2s}}$ for exponentially decaying φ. Again, one might think about Gamma processes and tempered stable processes. This should be compared to the rates of convergence which are found (an known to be minimax optimal) in density deconvolution problems. It should not come as a surprise, that we recover in the continuous limit (that is, for Δ close to zero) the rates which are known from density estimation with pointwise loss.

(ii) When longing to estimate $\mu([a,b]) = \int 1([a,b])(x)\mu(\,\mathrm{d}x)$ for some compact set $[a,b]$ bounded away from the origin, we have $s = 1$. The rate is parametric in the Compound Poisson case or for Gamma processes observed at a high enough frequency. Else, the rate is polynomial for polynomial decay and logarithmic for exponential decay of the characteristic function.

3 Adaptive estimation

3.1 The problem at hand

Let a collection $\mathcal{M} = \{m_1, \cdots, m_n\} \subseteq \mathbb{N}$ of indices be given and let $\mathcal{H} := \{h_1, \cdots, h_{m_n}\} := \{\frac{1}{m_1}, \cdots, \frac{1}{m_n}\}$ be a collection of bandwidths associated with $\mathcal{M}$.

For notational simplicity, we shall suppress in this section the dependence on Δ and assume that the distance between the observations of the Lévy process X is equal to one. Moreover, we slightly change the notation and write, when referring to the kernel estimator defined by (2.14), $\widehat{\theta}_{m,n}$ instead of $\widehat{\theta}_{1,h_m,n}$.

The goal of this section is to provide a strategy for the data driven choice of the smoothing index within the collection $\mathcal{M}$ and to derive, for the corresponding estimator $\widehat{\theta}_{\widehat{m},n}$, the oracle inequality

$$
\begin{aligned}
&\mathbb{E}\left[|\theta - \widehat{\theta}_{\widehat{m},n}|^2\right] \\
\leq\ & C \inf_{m\in\mathcal{M}} \left\{ |\theta - \theta_{m_n}|^2 + \sup_{\substack{k\geq m \\ k\in\mathcal{M}}} |\theta_k - \theta_m|^2 + \mathrm{pen}(m) \right\} + O\left(n^{-1}\right),
\end{aligned}
$$

with

$$
\theta_m := \frac{1}{2\pi}\int f(x)\left(\mathrm{K}_{\frac{1}{m}} * \mu\right)(x)\,\mathrm{d}x,
$$

with some constant C which does not depend on the unknown smoothness parameters and a penalty term $\mathrm{pen}(m)$ to be specified, which equals, up to some logarithmic factor, the quantity

$$\begin{aligned}\frac{1}{n}\sigma_m^2 \quad := \quad & \frac{n^{-1}}{2\pi^2}\Bigg\{ C_1 \int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right|^2 \left|\mathrm{K}\left(\frac{u}{m}\right)\right|^2 \mathrm{d}u \\ & \wedge C_2 \left(\int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right| \left|\mathrm{K}\left(\frac{u}{m}\right)\right| \mathrm{d}u \right)^2 \Bigg\}\end{aligned}$$

which bounds the error within the model.

3.2 Some heuristics

We start by giving some intuition on how the model selection procedure should work without going into the technical details. These considerations will be made precise in the next section.

If the characteristic function φ appearing in the denominator were feasible (which is, of course, not the case in the present setting), the way to go would be to estimate the quantities $|\theta_k - \theta_m|^2$ involved in the oracle bound by their corrected version, that is, to consider $|\widehat{\theta}_k - \widehat{\theta}_m|^2 - \mathrm{H}^2(m,k)$, with some deterministic correction term $\mathrm{H}^2(m,k)$ which is chosen large enough to ensure that with high probability,

$$|\widehat{\theta}_k - \widehat{\theta}_m|^2 - \mathrm{H}^2(m,k) \leq |\theta_k - \theta_m|^2 \ \forall m,k \in \mathcal{M}.$$

On the other hand, $\mathrm{H}^2(m,k)$ should ideally not be much larger than the variance term.

Typically, this would lead to choosing

$$\mathrm{H}^2(m,k) := \frac{1}{n}\rho\lambda_{m,k}^2 \left(\sigma_{m,k}^2 + x_{m,k}^2\right)$$

with some positive constant ρ to be appropriately chosen and

$$\begin{aligned}\sigma_{m,k}^2 \quad := \quad & \frac{1}{2\pi^2}\Bigg\{ C_1 \int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right|^2 \left|\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right|^2 \mathrm{d}u \\ & \wedge\ C_2 \left(\int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right| \left|\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right| \mathrm{d}u \right)^2 \Bigg\}\end{aligned}$$

and

$$x_{m,k} := \frac{1}{\sqrt{n}}\frac{1}{2\pi} \int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right| \left|\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right| \mathrm{d}u$$

and with logarithmic weights $\lambda_{m,k}$ chosen large enough to ensure that $\sum\limits_{\substack{k>m \\ k\in\mathcal{M}}} e^{-\lambda_{m,k}} < \infty$.

Indeed, this is the fundamental idea about model selection via penalization: Some *deterministic* term is applied in order to control the fluctuation of certain *stochastic* quantities, uniformly over some countable index set. For further reading, we refer to [2, 18, 3] among others.

Obviously, the situation is different in the present set up since the definition of the correction term $\mathrm{H}^2(m,k)$ involves the characteristic function in the denominator which is itself unknown.

It is well intuitive to replace the unknown characteristic function appearing in the denominator by its truncated empirical version $\frac{1}{\widetilde{\varphi}_n(u)} = \frac{1(\{|\widehat{\varphi}_n(u)|\geq n^{-1/2}\})}{\widehat{\varphi}_n(u)}$, thus considering

$$\widetilde{\sigma}^2_{m,k} \;:=\; \frac{1}{2\pi^2}\Bigg\{ C_1 \int \left|\frac{\mathcal{F}f(-u)}{\widetilde{\varphi}_n(u)}\right|^2 |\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)|^2 \,\mathrm{d}u \\ \wedge C_2 \left(\int \left|\frac{\mathcal{F}f(-u)}{\widetilde{\varphi}_n(u)}\right| |\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)|\,\mathrm{d}u\right)^2 \Bigg\}$$

instead of $\sigma^2_{m,k}$ and a stochastic version $\widetilde{x}_{m,k}$ of $x_{m,k}$ and to introduce a stochastic correction term $\widetilde{\mathrm{H}}^2(m,k) = \rho\widetilde{\lambda}^2_{m,k}\left(\widetilde{\sigma}^2_{m,k} + \widetilde{x}^2_{m,k}\right)$ rather than $\mathrm{H}^2(m,k)$.

Now, it is obvious that $\frac{1}{\widetilde{\varphi}_n(u)}$ may be sufficiently close to $\frac{1}{\varphi(u)}$ for large values of $|\varphi(u)|$, but is a drastic underestimate if $|\varphi(u)|$ is small. Consequently, the stochastic bias correction term $\widetilde{\mathrm{H}}^2(m,k)$ will systematically underestimate the true $\mathrm{H}^2(m,k)$, for which reason it seems doubtful if penalizing with $\widetilde{\mathrm{H}}^2(m,k)$ can possibly make sense.

In the setting of nonparametric estimation for Lévy processes with L^2-loss, Comte and Genon-Catalot [7] have dealt with the problem of the unknown variance by proposing an a priori assumption on the size of the collection $\mathcal{M}$ of smoothing parameters. However, this approach turns out to be critical since this assumption depends itself on the unknown decay of φ.

Only recently, Comte and Lacour [9] have proposed an approach towards model selection with unknown variance, which does not depend on any prior knowledge of the smoothness parameters . However, this approach is designed for L^2-loss and spectral cutoff estimation and the generalization to the estimation of linear functionals and general kernels is not straightforward. Moreover, it would lead to a loss of polynomial order in the present model. For this reason, we propose, in what follows, a different strategy.

Roughly speaking, the strategy in the above mentioned papers can be described as follows: At the first stage, one penalizes with some theoretical correction term which involves the unknown characteristic function. This makes the model selection procedure work as if φ were feasible. At the second stage, one has to control the fluctuation of the stochastic penalty round the theoretical penalty.

Compared to this, we undergo here some change of perspective by having a direct look at the stochastic penalty term:

For one thing, we may hope that, for large values of $|\varphi(u)|$, $\frac{1}{\widetilde{\varphi}_n(u)}$ is not only pointwise, but *uniformly* close to $\frac{1}{\varphi(u)}$, for which reason working with $\widetilde{\mathrm{H}}^2(m,k)$ rather than $\mathrm{H}^2(m,k)$ will work out right.

On the other hand, there remains the undeniable fact that for $|\varphi(u)|$ small, $\left|\frac{1}{\widetilde{\varphi}_n(u)}\right|$ is by no means close to $\left|\frac{1}{\varphi(u)}\right|$, but a systematic underestimate. For this reason penalizing with $\widetilde{\mathrm{H}}^2(m,k)$ rather than $\mathrm{H}^2(m,k)$ seems hopeless.

Still, one may ask oneself what is the use in penalizing at all. Certainly, the point about correcting with $\mathrm{H}^2(m,k)$ is that one wishes that with hight probability

$$|\widehat{\theta}_k - \widehat{\theta}_m|^2 - \mathrm{H}^2(m,k) \leq |\theta_k - \theta_m|^2 \, \forall m,k \in \mathcal{M}. \tag{3.1}$$

Now, if φ is unknown and has to be estimated, we must beware of the fact that the empirical version $\frac{1}{\widetilde{\varphi}_n}$ is involved not only in the definition of the stochastic correction term $\widetilde{\mathrm{H}}(m,k)$, but also appears in the definition of

$$\widehat{\theta}_k - \widehat{\theta}_m = \int \mathcal{F}f(-u) \frac{\widehat{\varphi}'_n(u)}{i\widetilde{\varphi}_n(u)} \left(\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right) \mathrm{d}u.$$

When considering small values of $|\varphi(u)|$ there is certainly no danger of overestimating in $|\widehat{\theta}_k - \widehat{\theta}_m|^2$. For this reason, subtracting some penalty term is simply not necessary at this stage, for which reason underestimating the quantity in $\mathrm{H}^2(m,k)$ as well causes no damage.

What remains to be done is to give some rigorous argument which allows to control the fluctuation of $\frac{1}{\widetilde{\varphi}_n(u)}$ round $\frac{1}{\varphi(u)}$ *uniformly* on the whole real line.

3.3 Adaptive estimation procedure and oracle bound

We have argued that we will need some result allowing to control the fluctuation of the empirical characteristic function in the denominator round its target uniformly on the whole real line. This will be done by applying concentration inequalities of Talagrand type.

For this purpose, we will need an alternative definition of an estimator of $\frac{1}{\varphi}$ and of the kernel estimator $\widehat{\theta}_{m,n}$:

3.1 Definition.

(i) Let the weight function w be defined by

$$w(u) = (\log(e + |u|))^{-\frac{1}{2}-\delta} \tag{3.2}$$

for some $\delta > 0$. For some constant κ to be specified, let

$$\widetilde{\widetilde{\varphi}}_n^{\delta,\kappa}(u) := \begin{cases} \widehat{\varphi}_n(u), & \text{if } |\widehat{\varphi}_n(u)| \geq \kappa(\log n)^{\frac{1}{2}} w(u)^{-1} n^{-\frac{1}{2}} \\ \kappa(\log n)^{\frac{1}{2}} w(u)^{-1} n^{-\frac{1}{2}}, & \text{else.} \end{cases} \tag{3.3}$$

The corresponding estimator of $\frac{1}{\varphi(u)}$ is $\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} := \frac{1}{\widetilde{\widetilde{\varphi}}_n^{\delta,\kappa}(u)}$.

(ii) In what follows, let $\widehat{\theta}_{m,n} := \widehat{\theta}_{1,\frac{1}{m},n}$ be defined as in Definition 2.3, apart from the fact that $\frac{1}{\widetilde{\varphi}_{1,n}}$ is replaced in (2.14) by $\frac{1}{\widetilde{\widetilde{\varphi}}_n}$, defined as in (i).

What will be important about this redefinition is the fact that we have introduced an extra logarithmic factor which will enable us to give uniform control on $\left|\frac{1}{\varphi} - \frac{1}{\widetilde{\widetilde{\varphi}}_n(u)}\right|^2$. More precisely, we can proof the following key result which makes the well known Lemma by Neumann (see Lemma 2.1 in [20]) uniform on the real line:

3.2 Lemma. *Let c_1 be the constant appearing in Lemma 5.4. Let $\frac{1}{\widetilde{\widetilde{\varphi}}_n}$ be defined by Definition 3.1 (i) with κ be chosen such that for some $\gamma > 0$, we have $\kappa \geq 2(\sqrt{2c_1} + \gamma)$. Then we have for some constant $C_{N,K}$ depending on the choice of κ, γ and δ:*

$$\mathbb{E}\left[\sup_{u \in \mathbb{R}} \frac{\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2}{\frac{(\log n) w(u)^{-2} n^{-1}}{|\varphi(u)|^4} \wedge \frac{1}{|\varphi(u)|^2}}\right] \leq C_{N,K}$$

First of all, we observe that thanks to Lemma 3.2, for the squared risk of the newly defined estimator $\widehat{\theta}_{m,n}$, we have

$$\mathbb{E}\left[\left|\theta - \widehat{\theta}_{m,n}\right|^2\right] \leq |\theta - \theta_m|^2 + \frac{1}{n}\sigma_{m,w}^2$$

with

$$\begin{aligned} \sigma_{m,w}^2 \quad := \quad & \frac{\log n}{2\pi^2}\Bigg\{C_1 \int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right|^2 \left|\mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right|^2 w(u)^{-2}\,\mathrm{d}u \\ & \wedge\, C_2\left(\int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right| \left|\mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right| w(u)^{-1}\,\mathrm{d}u\right)^2\Bigg\}, \end{aligned}$$

that is, the upper risk bound is preserved up to a logarithmic factor. The proof is the same as the proof of the upper risk bound given in Theorem 2.4.

Let us introduce some definitions which will be needed in the sequel.

For $m, k \in \mathcal{M}$, let

$$\begin{aligned}\sigma^2_{m,k,w} \quad := \quad & \frac{\log n}{2\pi^2}\Bigg\{\bar{C}_1 \int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right|^2 \left|\mathcal{F}\,\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\,\mathrm{K}\left(\frac{u}{m}\right)\right|^2 w(u)^{-2}\,\mathrm{d}u \\ & \wedge \bar{C}_2 \left(\int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right| \left|\mathcal{F}\,\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\,\mathrm{K}\left(\frac{u}{m}\right)\right| w(u)^{-1}\,\mathrm{d}u\right)^2\Bigg\}.\end{aligned}$$

Let

$$x_{m,k,w} := \frac{\log n}{\sqrt{n}}\frac{1}{2\pi}\int \left|\frac{\mathcal{F}f(-u)}{\varphi(u)}\right| \left|\mathcal{F}\,\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\,\mathrm{K}\left(\frac{u}{m}\right)\right| w(u)^{-1}\,\mathrm{d}u$$

and

$$\begin{aligned}\lambda_{m,k,w} \quad := \quad & \left(\log n + \log\left(x^2_{m,k,w}(k-m)^2\right)\right) \\ & \left\{\log\left(\left(\sigma^2_{m,k,w} + x^2_{m,k,w}\right)(k-m)^2\right) + \log\left(\log n + \log\left(x^2_{m,k,w}(k-m)^2\right)\right)\right\}\end{aligned}$$

Finally, let

$$\mathrm{H}^2(m,k) \quad := \quad \rho\frac{1}{n}\lambda^2_{m,k,w}\left(\sigma^2_{m,k,w} + x^2_{m,k,w}\right)$$

and

$$\mathrm{pen}(m) := \mathrm{H}^2(0,m).$$

The stochastic counterparts $\widetilde{\sigma}^2_{m,k,w}$, $\widetilde{x}_{m,k,w}$, $\widetilde{\lambda}_{m,k,w}$, $\widetilde{\mathrm{H}}^2(m,k)$ and $\widetilde{\mathrm{pen}}(m)$ are defined by replacing, in each of the above definitions, $\frac{1}{\varphi}$ by $\frac{1}{\widetilde{\varphi}_n}$.

Now, let the random smoothing parameter be defined to be

$$\widehat{m} := \operatorname*{arginf}_{m\in\mathcal{M}}\left\{\sup_{\substack{k>m\\k\in\mathcal{M}}}\left\{\left|\widehat{\theta}_k - \widehat{\theta}_m\right|^2 - \widetilde{\mathrm{H}}^2(m,k)\right\} + \widetilde{\mathrm{pen}}(m)\right\}. \tag{3.4}$$

We are now ready to formulate the main result of this section:

3.3 Theorem. *Let observations $X_1, \cdots, X_{2n}$ of a Lévy process be given. Let $\mathcal{M} = \{1, \cdots, m_n\}$. Assume that for some positive constant η, $\mathbb{E}\left[\exp\left(\eta X_1\right)\right] < \infty$. Assume, moreover, that $C_1 \leq \bar{C}_1$ and $C_2 \leq \bar{C}_2$.*

For $m \in \mathcal{M}$, let $\widehat{\theta}_{m,n}$ be defined by Definition 3.1. Let $\widehat{m}$ be defined by (3.4) and assume that we have and $\rho \geq 128\frac{\sqrt{2}}{3}$ and $\kappa \geq 2(\sqrt{2c_1} + \gamma)$. Then we can estimate

$$\begin{aligned} & \mathbb{E}\left[\left|\theta - \widehat{\theta}_m\right|^2\right] \\ \leq \quad & C \inf_{m\in\mathcal{M}}\left\{|\theta - \theta_{m_n}|^2 + \sup_{\substack{k>m\\k\in\mathcal{M}}}|\theta_k - \theta_m|^2 + \mathrm{pen}(m)\right\} + O\left(n^{-1}\right)\end{aligned}$$

for some positive constant C which does not depend on the decay of φ nor on the smoothness of μ.

Theorem 3.3 will tell us that the estimation procedure attains, up to a logarithmic loss, the optimal rates of convergence. It is worth mentioning that we can relax the exponential moment condition on X_1, but at the cost of losing an polynomial factor.

Our reasoning is not particular to the Lévy model nor to the estimation of linear functionals, but generalizes to the setting of nonparametric deconvolution with unknown error distribution and to L^2-loss. A detailed discussion on the subject will be given in [17]

4 Acknowledgement

I want to thank Fabienne Comte and Valentine Genon-Catalot for giving me the opportunity for a research stay at Université Paris Déscartes and for inspiring discussions on model selection. Moreover, I thank Markus Reiß for useful hints and comments and Matthias Trabs for careful reading.

5 Proofs

5.1 Proofs of the main results of Section 2

The following lemma is the key result for the proofs Theorem 2.4 and Theorem 2.5.

5.1 Lemma. *Let $\widehat{\varphi}'_{\Delta,n}$ and $\frac{1}{\widetilde{\varphi}_{\Delta,n}}$ and be defined by (2.11) and (2.12). Then we can estimate*

$$\left|\mathbb{E}\left[\left(\frac{\frac{1}{\Delta}\widehat{\varphi}'_{\Delta,n}(u)}{\widetilde{\varphi}_{\Delta,n}(u)}-\frac{\frac{1}{\Delta}\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\left(\frac{\frac{1}{\Delta}\widehat{\varphi}'_{\Delta,n}(-v)}{\widetilde{\varphi}_{\Delta,n}(-v)}-\frac{\frac{1}{\Delta}\varphi'_{\Delta}(-v)}{\varphi_{\Delta}(-v)}\right)\right]\right|$$
$$\leq C\left(\frac{T^{-1}}{|\varphi_{\Delta}(u)||\varphi_{\Delta}(-v)|}\wedge 1\right)\big(|\Psi''(u-v)|+|\Psi'(u-v)|+|\Psi'(u)||\Psi'(-v)|\big)$$

with some universal constant C.

Proof. We start by noting that for some constant $C_{N,k}$, we have

$$\mathbb{E}\left[\left|\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)}-\frac{1}{\varphi_{\Delta}(u)}\right|^k\right]\leq C_{N,k}\left(\frac{T^{-\frac{k}{2}}}{|\varphi_{\Delta}(u)|^{2k}}\wedge\frac{1}{|\varphi_{\Delta}(u)|^k}\right), \tag{5.1}$$

which is a direct consequence of Neumann's Lemma, drawing back the dependence on Δ.

We can write

$$\mathbb{E}\left[\left(\frac{\widehat{\varphi}'_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\left(\frac{\widehat{\varphi}'_{\Delta_n}(-v)}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{\varphi'_\Delta(-v)}{\varphi_\Delta(-v)}\right)\right] \tag{5.2}$$

$$= \mathbb{E}\left[\left(\frac{(\widehat{\varphi}'_{\Delta_n}(u) - \varphi'_\Delta(u))}{\widetilde{\varphi}_{\Delta_n}(u)} + \varphi'_\Delta(u)\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{1}{\varphi_\Delta(u)}\right)\right)\right. \tag{5.3}$$

$$\left.\left(\frac{(\widehat{\varphi}'_{\Delta_n}(-v) - \varphi'_\Delta(-v))}{\widetilde{\varphi}_{\Delta_n}(-v)} + \varphi'_\Delta(-v)\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{1}{\varphi_\Delta(-v)}\right)\right)\right] \tag{5.4}$$

Using the fact that $\widehat{\varphi}'_{\Delta_n}$ and $\frac{1}{\widetilde{\varphi}_{\Delta_n}}$ are independent by construction and that $\widehat{\varphi}'_{\Delta_n}(u) - \varphi'_\Delta(u)$ and $\widehat{\varphi}'_\Delta(-v) - \varphi'_\Delta(-v)$ are centered, we find that

$$\mathbb{E}\left[\left(\frac{(\widehat{\varphi}'_{\Delta_n}(u) - \varphi'_\Delta(u))}{\widetilde{\varphi}_{\Delta_n}(u)} + \varphi'_\Delta(u)\left(\frac{1}{\widetilde{\varphi}_{\Delta_{n(u)}}} - \frac{1}{\varphi_\Delta(u)}\right)\right)\right. \tag{5.5}$$

$$\left.\left(\frac{(\widehat{\varphi}'_{\Delta_n}(-v) - \varphi'_\Delta(-v))}{\widetilde{\varphi}_{\Delta_n}(-v)} + \varphi'_\Delta(-v)\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{1}{\varphi_\Delta(-v)}\right)\right)\right] \tag{5.6}$$

$$= \mathbb{E}\left[\left(\widehat{\varphi}'_{\Delta_n}(u) - \varphi'_\Delta(u)\right)\left(\widehat{\varphi}'_{\Delta_n}(-v) - \varphi'_\Delta(-v)\right)\right]\mathbb{E}\left[\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)\widetilde{\varphi}_{\Delta_n}(-v)}\right] \tag{5.7}$$

$$+ \quad \varphi'_\Delta(u)\varphi'_\Delta(-v)\,\mathbb{E}\left[\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{1}{\varphi_\Delta(u)}\right)\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{1}{\varphi_\Delta(-v)}\right)\right] \tag{5.8}$$

$$= \operatorname{Cov}\left(\widehat{\varphi}'_{\Delta_n}(u), \widehat{\varphi}'_{\Delta_n}(-v)\right)\mathbb{E}\left[\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)\widetilde{\varphi}_{\Delta_n}(-v)}\right] \tag{5.9}$$

$$+ \quad \varphi'_\Delta(u)\varphi'_\Delta(-v)\,\mathbb{E}\left[\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{1}{\varphi_\Delta(u)}\right)\left(\frac{1}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{1}{\varphi - \Delta(-v)}\right)\right] \tag{5.10}$$

$$=: \operatorname{Cov}\left(\widehat{\varphi}'_{\Delta_n}(u), \widehat{\varphi}'_{\Delta_n}(-v)\right)\mathbb{E}\left[\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)\widetilde{\varphi}_{\Delta_n}(-v)}\right] \tag{5.11}$$

$$+ \quad \varphi'_\Delta(u)\varphi'_\Delta(-v)\,\mathbb{E}\left[\mathrm{R}_{\Delta_n}(u)\mathrm{R}_{\Delta_n}(-v)\right]. \tag{5.12}$$

The Cauchy-Schwarz-inequality and then an application of (5.1) imply

$$\mathbb{E}\left[|\mathrm{R}_{\Delta_n}(u)\mathrm{R}_{\Delta_n}(-v)|\right] \tag{5.13}$$

$$\leq \left(\mathbb{E}\left[|\mathrm{R}_{\Delta_n}(u)|^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[|\mathrm{R}_{\Delta_n}(-v)|^2\right]\right)^{\frac{1}{2}} \tag{5.14}$$

$$\leq C_{N,2}\left(\frac{T^{-1}}{|\varphi_\Delta(u)|^2|\varphi_\Delta(-v)|^2} \wedge \frac{1}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}\right). \tag{5.15}$$

Next, using the triangular inequality, again (5.1) and then the same reasoning as in (5.13)-(5.15), we find that

$$
\begin{aligned}
& \mathbb{E}\left[\left|\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)\widetilde{\varphi}_{\Delta_n}(-v)}\right|\right] && (5.16)\\
\leq\ & \frac{1}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}+\frac{1}{|\varphi_\Delta(u)|}\,\mathbb{E}\left[|\mathrm{R}_{\Delta_n}(-v)\right] && (5.17)\\
+\ & \frac{1}{|\varphi_\Delta(-v)|}\,\mathbb{E}\left[|\mathrm{R}_{\Delta_n}(u)|\right]+\mathbb{E}\left[|\mathrm{R}_{\Delta_n}(u)|\,|\mathrm{R}_{\Delta_n}(-v)|\right] && (5.18)\\
\leq\ & (1+2C_{N,1}+C_{N,2})\frac{1}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}. && (5.19)
\end{aligned}
$$

Moreover, by definition of $\frac{1}{\widetilde{\varphi}_n}$, we have

$$
\mathbb{E}\left[\left|\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)\widetilde{\varphi}_{\Delta_n}(-v)}\right|\right]\leq T. \tag{5.20}
$$

Next, we calculate

$$
\begin{aligned}
& \mathrm{Cov}(\widehat{\varphi}'_{\Delta_n}(u),\widehat{\varphi}'_{\Delta_n}(v)) && (5.21)\\
=\ & n^{-1}\left(\mathbb{E}\left[(iZ_\Delta)^2e^{i(u-v)Z_\Delta}\right]-\mathbb{E}\left[iZ_\Delta e^{iuZ_\Delta}\right]\mathbb{E}\left[iZ_\Delta e^{-ivZ_\Delta}\right]\right) && (5.22)\\
=\ & n^{-1}\left(\varphi''_\Delta(u-v)-\varphi'_\Delta(u)\varphi'_\Delta(-v)\right) && (5.23)
\end{aligned}
$$

Moreover, we clearly have

$$
\begin{aligned}
\left|\varphi'_\Delta(u)\right|\left|\varphi'_\Delta(-v)\right| &= \left|\Delta\Psi'(u)\varphi_\Delta(u)\right|\left|\Delta\Psi'(-v)\varphi_\Delta(-v)\right|\\
&\leq |\Delta\Psi'(u)||\Delta\Psi'(-v)|.
\end{aligned}
$$

and

$$
\begin{aligned}
|\varphi''_\Delta(u-v)| &= \left|\Delta\Psi''(u-v)\varphi_\Delta(u-v)+\Delta^2(\Psi'(u-v))^2\varphi_\Delta(u-v)\right|\\
&\leq \Delta|\Psi''(u-v)|+\Delta^2(\Psi'(u-v))^2
\end{aligned}
$$

Putting (5.16)-(5.19), (5.20) and (5.21)-(5.23) together, the expression appearing in (5.11) can be estimated as follows:

$$
\begin{aligned}
& \left|\mathrm{Cov}\left(\widehat{\varphi}'_{\Delta_n}(u),\widehat{\varphi}'_{\Delta_n}(v)\right)\right|\left|\mathbb{E}\left[\frac{1}{\widetilde{\varphi}_{\Delta_n}(u)\widetilde{\varphi}_{\Delta_n}(-v)}\right]\right| && (5.24)\\
\leq\ & (1+2C_{N,1}+C_{N,2})\left(T\wedge\frac{1}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}\right) && (5.25)\\
& n^{-1}\Big(\left|\Delta\Psi''(u-v)\right|+|\Delta\Psi'(u-v)|^2+\left|\Delta\Psi'(u)\right|\left|\Delta\Psi'(-v)\right|\Big) && (5.26)\\
\leq\ & (1+2C_{N,1}+C_{N,2})\Delta^2\left(1\wedge\frac{T^{-1}}{|\varphi(u)||\varphi(-v)|}\right) && (5.27)\\
& \Big(\left|\Psi''(u-v)\right|+\Delta|\Psi'(u-v)|^2+\Delta\left|\Psi'(u)\right|\left|\Psi'(-v)\right|\Big). && (5.28)
\end{aligned}
$$

Next, using (5.13)-(5.15) and then the fact that $\varphi'_\Delta(u) = \Delta\Psi'(u)\varphi_\Delta(u)$, the expression in (5.12) and be estimated:

$$
\begin{aligned}
& |\varphi'_\Delta(u)||\varphi'_\Delta(-v)|\,|\mathbb{E}\left[\mathrm{R}_{\Delta_n}(u)\mathrm{R}_{\Delta_n}(-v)\right]| && (5.29)\\
\leq\ & C_{N,2}\left(\frac{T^{-1}}{|\varphi_\Delta(u)|^2|\varphi_\Delta(-v)|^2} \wedge \frac{1}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}\right)\left|\varphi'_\Delta(u)\right|\left|\varphi'_\Delta(-v)\right| && (5.30)\\
=\ & C_{N,2}\left(\frac{T^{-1}}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|} \wedge 1\right)\Delta^2\left|\Psi'(u)\right|\left|\Psi'(-v)\right|. && (5.31)
\end{aligned}
$$

Putting (5.24)-(5.28) and (5.29)-(5.31) together, we have shown that

$$
\begin{aligned}
& \frac{1}{\Delta^2}\left|\mathbb{E}\left[\left(\frac{\widehat{\varphi}'_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\left(\frac{\widehat{\varphi}'_{\Delta_n}(-v)}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{\varphi'_\Delta(-v)}{\varphi_\Delta(-v)}\right)\right]\right|\\
\leq\ & C\left(\frac{T^{-1}}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|} \wedge 1\right)\left(\left|\Psi''(u-v)\right| + |\Psi'(u-v)|^2|\Psi'(u)||\Psi'(-v)|\right),
\end{aligned}
$$

which is the statement of the lemma. $\square$

We can now use Lemma 5.1 to prove Theorem 2.4.

Proof of Theorem 2.4. The risk of $\widehat{\theta}_{h,n}$ can be decomposed as follows: With $\theta_h := \frac{1}{2\pi}\int f(x)(\mathrm{K}_h * g)(x)\,\mathrm{d}x$, we have

$$
\begin{aligned}
& \mathbb{E}\left[\left|\theta - \widehat{\theta}_{h,n}\right|^2\right]\\
\leq\ & 2\left|\theta - \theta_h\right|^2 + 2\,\mathbb{E}\left[\left|\theta_h - \widehat{\theta}_{h,n}\right|^2\right]\\
=\ & 2\left|\int f(x)\mu(\,\mathrm{d}x) - \int f(x)(\mathrm{K}_h * \mu)(x)\,\mathrm{d}x\right|^2\\
+\ & 2\,\mathbb{E}\left[\left|\int f(x)(\mathrm{K}_h * \mu)(x)\,\mathrm{d}x - \frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{\frac{1}{\Delta}\widehat{\varphi}_{\Delta_n}(u)}{i\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right].
\end{aligned}
$$

By assumption on K, we have $\mathcal{F}K_h\mathcal{F}f \in \mathrm{L}^1(\mathbb{R})$, so we can pass to the Fourier domain and find that

$$
\begin{aligned}
& \mathbb{E}\left[\left|\int f(x)(\mathrm{K}_h * \mu)(x)\,\mathrm{d}x - \frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{\widehat{\varphi}_{\Delta_n}(u)}{i\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right]\\
=\ & \mathbb{E}\left[\left|\frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\mathcal{F}\mu(u)\,\mathrm{d}u - \frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{\frac{1}{\Delta}\widehat{\varphi}_{\Delta_n}(u)}{i\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right]\\
=\ & \mathbb{E}\left[\left|\frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{1}{\Delta}\left(\frac{\widehat{\varphi}'_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\,\mathrm{d}u\right|^2\right].
\end{aligned}
$$

An application of Fubini's theorem yields

$$
\begin{aligned}
&\mathbb{E}\left[\left|\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{1}{\Delta}\left(\frac{\widehat{\varphi}'_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)}-\frac{\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\,\mathrm{d}u\right|^2\right]\\
=\;&\int\int \mathcal{F}f(-u)\mathcal{F}f(v)\mathcal{F}\,\mathrm{K}_h(u)\mathcal{F}\,\mathrm{K}_h(-v)\\
&\qquad\times\frac{1}{\Delta^2}\,\mathbb{E}\left[\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)}-\frac{\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\left(\frac{\widehat{\varphi}_{\Delta_n}(-v)}{\widetilde{\varphi}_{\Delta_n}(-v)}-\frac{\varphi'_{\Delta}(-v)}{\varphi_{\Delta}(-v)}\right)\right]\mathrm{d}u\,\mathrm{d}v
\end{aligned}
$$

and next, Lemma 5.1 gives

$$
\begin{aligned}
&\int\int \mathcal{F}f(-u)\mathcal{F}f(v)\mathcal{F}\,\mathrm{K}_h(u)\mathcal{F}\,\mathrm{K}_h(-v)\\
&\qquad\times\frac{1}{\Delta^2}\,\mathbb{E}\left[\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)}-\frac{\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\left(\frac{\widehat{\varphi}_{\Delta_n}(-v)}{\widetilde{\varphi}_{\Delta_n}(-v)}-\frac{\varphi'_{\Delta}(-v)}{\varphi_{\Delta}(-v)}\right)\right]\mathrm{d}u\,\mathrm{d}v\\
\leq\;& CT^{-1}\Bigg\{\int\int\frac{|\mathcal{F}f(-u)||\mathcal{F}f(v)|}{|\varphi_{\Delta}(u)||\varphi_{\Delta}(-v)|}\left|\Psi'(u)\right|\left|\Psi'(-v)\right|\,\mathrm{d}u\,\mathrm{d}v\\
+\;&\int\int\frac{|\mathcal{F}f(-u)|\,|\mathcal{F}f(v)|}{|\varphi_{\Delta}(u)||\varphi_{\Delta}(-v)|}\,|\mathrm{K}_h(u)|\,|\mathrm{K}_h(-v)|\left(\left|\Psi''(u-v)\right|+\left|\Psi'(u-v)\right|\right)\,\mathrm{d}u\,\mathrm{d}v\Bigg\}.
\end{aligned}
$$

In case that $\Psi''\in \mathrm{L}^1(\mathbb{R})$ and $\Psi'\in\mathrm{L}^2(\mathbb{R})$, we apply the Cauchy-Schwarz inequality and Fubini's theorem to find that

$$
\begin{aligned}
&\int\int\frac{|\mathcal{F}f(-u)|\,|\mathcal{F}f(v)|}{|\varphi_{\Delta}(u)||\varphi_{\Delta}(-v)|}\,|\mathrm{K}_h(u)|\,|\mathrm{K}_h(-v)|\left(\left|\Psi''(u-v)\right|+\left|\Psi'(u-v)\right|^2\right)\,\mathrm{d}u\,\mathrm{d}v\\
\leq\;&\int\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_{\Delta}(u)|^2}\,|\mathcal{F}\,\mathrm{K}_h(u)|^2\,(\left|\Psi''(u-v)\right|+|\Psi'(u-v)|^2)\,\mathrm{d}u\,\mathrm{d}v\\
=\;&\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_{\Delta}(u)|^2}\,|\mathcal{F}\,\mathrm{K}_h(u)|^2\int(|\Psi''(u-v)|+|\Psi'(u-v)|^2)\,\mathrm{d}v\,\mathrm{d}u\\
\leq\;&\sup_{u\in\mathbb{R}}\int(|\Psi''_{\Delta}(u-v)|+|\Psi'(u-v)|^2)\,\mathrm{d}v\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_{\Delta}(u)|^2}\,|\mathcal{F}\,\mathrm{K}_h(u)|^2\,\mathrm{d}u\\
\leq\;&\left(\int(|\Psi''(x)|\,\mathrm{d}x+\int|\Psi'(x)|^2\,\mathrm{d}x\right)\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_{\Delta}(u)|^2}\,|\mathcal{F}\,\mathrm{K}_h(u)|^2\,\mathrm{d}u.
\end{aligned}
$$

Another application of the Chauchy-Schwarz-inequality gives

$$
\begin{aligned}
&\int\int\frac{|\mathcal{F}f(-u)||\mathcal{F}f(v)|}{|\varphi_{\Delta}(u)||\varphi_{\Delta}(-v)|}\,|\mathrm{K}_h(u)|\,|\mathrm{K}_h(-v)|\,|\Psi'(u)||\Psi'(-v)|\,\mathrm{d}u\,\mathrm{d}v\\
=\;&\left(\int\frac{|\mathcal{F}f(-u)|}{|\varphi_{\Delta}(u)|}|\mathcal{F}\,\mathrm{K}_h(u)||\Psi'(u)|\,\mathrm{d}u\right)^2\\
\leq\;&\int|\Psi'(u)|^2\,\mathrm{d}u\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_{\Delta}(u)|^2}|\mathcal{F}\,\mathrm{K}_h(u)|^2\,\mathrm{d}u.
\end{aligned}
$$

We have thus shown that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\frac{1}{2\pi}\int f(x)(\mathrm{K}_h*\mu)(x)\,\mathrm{d}x-\frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{\frac{1}{\Delta}\widehat{\varphi}_{\Delta_n}(u)}{i\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right]\\
\leq\ &\frac{C}{(2\pi)^2}T^{-1}\left(\int|\Psi''(x)|\,\mathrm{d}x+2\int|\Psi'(x)|^2\,\mathrm{d}x\right)\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_\Delta(u)|^2}|\mathcal{F}\,\mathrm{K}_h(u)|^2\,\mathrm{d}u.
\end{aligned}
$$

On the other hand, we can always estimate

$$
\begin{aligned}
&\int\int\frac{|\mathcal{F}f(-u)|\,|\mathcal{F}f(v)|}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}\,|\mathrm{K}_h(u)|\,|\mathrm{K}_h(-v)|\,\left(\left|\Psi''(u-v)\right|+|\Psi'(u-v)|^2\right)\mathrm{d}u\,\mathrm{d}v\\
\leq\ &\sup_{u,v\in\mathbb{R}}\left(\left|\Psi''(u-v)\right|+|\Psi'(u-v)|^2\right)\int\int\frac{|\mathcal{F}f(-u)|\,|\mathcal{F}f(v)|}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}\,|\mathrm{K}_h(u)|\,|\mathrm{K}_h(-v)|\,\mathrm{d}u\,\mathrm{d}v\\
\leq\ &\sup_{x\in\mathbb{R}}\left(\left|\Psi''(x)\right|+|\Psi'(x)|^2\right)\left(\int\frac{|\mathcal{F}f(-u)|}{|\varphi_\Delta(u)|}|\mathcal{F}\,\mathrm{K}_h(u)|\,\mathrm{d}u\right)^2
\end{aligned}
$$

and

$$
\begin{aligned}
&\int\int\frac{|\mathcal{F}f(-u)||\mathcal{F}f(v)|}{|\varphi_\Delta(u)||\varphi_\Delta(-v)|}|\mathcal{F}\,\mathrm{K}_h(u)||\mathcal{F}\,\mathrm{K}_h(-v)||\Psi'(u)||\Psi'(-v)|\,\mathrm{d}u\,\mathrm{d}v\\
=\ &\left(\int\frac{|\mathcal{F}f(-u)|}{|\varphi_\Delta(u)|}|\mathcal{F}\,\mathrm{K}_h(u)||\Psi'(u)|\,\mathrm{d}u\right)^2\\
\leq\ &\sup_{x\in\mathbb{R}}\left|\Psi'(x)\right|^2\left(\int\frac{|\mathcal{F}f(-u)|}{|\varphi_\Delta(u)|}|\mathcal{F}\,\mathrm{K}_h(u)|\,\mathrm{d}u\right)^2.
\end{aligned}
$$

This yields

$$
\begin{aligned}
&\mathbb{E}\left[\left|\frac{1}{2\pi}\int f(x)(\mathrm{K}_h*\mu)(x)\,\mathrm{d}x-\frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\,\mathrm{K}_h(u)\frac{\frac{1}{\Delta}\widehat{\varphi}_{\Delta_n}(u)}{i\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right]\\
\leq\ &\frac{C}{(2\pi)^2}T^{-1}\left(\sup_{x\in\mathbb{R}}|\Psi''(x)|+2\sup_{x\in\mathbb{R}}|\Psi'(x)|^2\right)\left(\int\frac{|\mathcal{F}f(-u)|}{|\varphi_\Delta(u)|}|\mathcal{F}\,\mathrm{K}_h(u)|\,\mathrm{d}u\right)^2.
\end{aligned}
$$

Putting the above results together, we have shown that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\theta-\widehat{\theta}_{h,n}\right|^2\right]\\
\leq\ &\left|\int f(x)\mu(\,\mathrm{d}x)-\int f(x)(\mathrm{K}_h*\mu)(x)\,\mathrm{d}x\right|^2\\
+\ &\frac{C}{(2\pi)^2}T^{-1}\Bigg\{\left(\int\left|\Psi''(x)\right|\,\mathrm{d}x+2\int|\Psi'(x)|^2\,\mathrm{d}x\right)\int\frac{|\mathcal{F}f(-u)|^2}{|\varphi_\Delta(u)|^2}|\mathcal{F}\,\mathrm{K}_h(u)|^2\,\mathrm{d}u\\
&\qquad\wedge\left(\sup_{x\in\mathbb{R}}|\Psi''(x)|+2\sup_{x\in\mathbb{R}}|\Psi'(x)|^2\right)\left(\int\frac{|\mathcal{F}f(-u)|}{|\varphi_\Delta(u)|}|\mathcal{F}\,\mathrm{K}_h(u)|\,\mathrm{d}u\right)^2\Bigg\},
\end{aligned}
$$

which is the statement of the theorem. □

Next, we prove Theorem 2.5.

Proof of Theorem 2.5. First, recall that we now assume that $|\mathcal{F}f| \in \mathrm{L}^1(\mathbb{R})$, so we certainly have $|\mathcal{F}f\mathcal{F}\mu| \leq \|\mathcal{F}\mu\|_\infty |\mathcal{F}f| \in \mathrm{L}^1(\mathbb{R})$. We can thus express θ in the Fourier domain, writing

$$\theta = \int f(x)\mu(\,\mathrm{d}x) = \frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\mu(u)\,\mathrm{d}u.$$

and express the squared risk of $\widehat{\theta}_n$ as follows:

$$\begin{aligned}
&\mathbb{E}\left[\left|\theta - \widehat{\theta}_n\right|^2\right] \\
=\;& \mathbb{E}\left[\left|\int f(x)\mu(\,\mathrm{d}x) - \frac{1}{2\pi}\int \mathcal{F}f(-u)\frac{\frac{1}{\Delta}\widehat{\varphi}'_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right] \\
=\;& \mathbb{E}\left[\left|\frac{1}{2\pi}\int \mathcal{F}f(-u)\mathcal{F}\mu(u)\,\mathrm{d}u - \frac{1}{2\pi}\int \mathcal{F}f(-u)\frac{\frac{1}{\Delta}\widehat{\varphi}'_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)}\,\mathrm{d}u\right|^2\right] \\
=\;& \frac{1}{\Delta^2}\,\mathbb{E}\left[\left|\frac{1}{2\pi}\int \mathcal{F}f(-u)\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\,\mathrm{d}u\right|^2\right].
\end{aligned}$$

Next, we can write for arbitrary $m \geq 0$:

$$\begin{aligned}
&\mathbb{E}\left[\left|\frac{1}{2\pi}\int \mathcal{F}f(-u)\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\,\mathrm{d}u\right|^2\right] \\
\leq\;& 2\,\mathbb{E}\left[\left|\frac{1}{2\pi}\int\limits_{\{|u|>\pi m\}} \mathcal{F}f(-u)\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\,\mathrm{d}u\right|^2\right] \\
+\;& 2\,\mathbb{E}\left[\left|\frac{1}{2\pi}\int\limits_{\{|u|\leq\pi m\}} \mathcal{F}f(-u)\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\,\mathrm{d}u\right|^2\right].
\end{aligned}$$

Applying Fubini's theorem, Lemma 5.1 and again the Cauchy-Schwarz inequality yields

$$\begin{aligned}
&\frac{1}{\Delta^2}\,\mathbb{E}\left[\left|\int\limits_{\{|u|>\pi m\}} \mathcal{F}f(-u)\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_\Delta(u)}{\varphi_\Delta(u)}\right)\,\mathrm{d}u\right|^2\right] \\
=\;& \int\limits_{\{|u|>\pi m\}}\int\limits_{\{|v|>\pi m\}} \mathcal{F}f(-u)\mathcal{F}f(v)
\end{aligned}$$

$$\times \frac{1}{\Delta^2}\,\mathbb{E}\left[\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\left(\frac{\widehat{\varphi}_{\Delta_n}(-v)}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{\varphi'_{\Delta}(-v)}{\varphi_{\Delta}(-v)}\right)\right] \mathrm{d}u\,\mathrm{d}v$$

$$\leq \int_{\{|u|>\pi m\}} \int_{\{|v|>\pi m\}} |\mathcal{F}f(-u)||\mathcal{F}f(v)|$$

$$\times \left(\left|\Psi''(u-v)\right| + |\Psi'(u-v)|^2 + \left|\Psi'(u)\right|\left|\Psi'(-v)\right|\right) \mathrm{d}u\,\mathrm{d}v$$

$$\leq \left(\int |\Psi''(x)|\,\mathrm{d}x + 2\int |\Psi'(x)|^2\,\mathrm{d}x\right) \int_{\{|u|>\pi m\}} |\mathcal{F}f(-u)|^2\,\mathrm{d}u$$

$$\wedge \left(\sup_{x\in\mathbb{R}} |\Psi''(x)| + 2\sup_{x\in\mathbb{R}} |\Psi'(x)|^2\right)\left(\int_{\{|u|>\pi m\}} |\mathcal{F}f(-u)|\,\mathrm{d}u\right)^2$$

Arguing along the same lines, we find that

$$\frac{1}{\Delta^2}\,\mathbb{E}\left[\left|\int_{\{|u|\leq\pi m\}} \mathcal{F}f(-u)\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\mathrm{d}u\right|^2\right]$$

$$= \int_{\{|u|\leq\pi m\}} \int_{\{|v|\leq\pi m\}} \mathcal{F}f(-u)\mathcal{F}f(v)$$

$$\times \frac{1}{\Delta^2}\,\mathbb{E}\left[\left(\frac{\widehat{\varphi}_{\Delta_n}(u)}{\widetilde{\varphi}_{\Delta_n}(u)} - \frac{\varphi'_{\Delta}(u)}{\varphi_{\Delta}(u)}\right)\left(\frac{\widehat{\varphi}_{\Delta_n}(-v)}{\widetilde{\varphi}_{\Delta_n}(-v)} - \frac{\varphi'_{\Delta}(-v)}{\varphi_{\Delta}(-v)}\right)\right] \mathrm{d}u\,\mathrm{d}v$$

$$\leq CT^{-1}\left\{\left(\int |\Psi''(x)|\,\mathrm{d}x + 2\int |\Psi'(x)|^2\,\mathrm{d}x\right) \int_{\{|u|\leq\pi m\}} \left|\frac{\mathcal{F}f(-u)}{\varphi_{\Delta}(u)}\right|^2 \mathrm{d}u\right.$$

$$\left.\left(\sup_{x\in\mathbb{R}} |\Psi''(x)| + 2\sup_{x\in\mathbb{R}} |\Psi'(x)|^2\right)\left(\int_{\{|u|\leq\pi m\}} \left|\frac{\mathcal{F}f(-u)}{\varphi_{\Delta}(u)}\right| \mathrm{d}u\right)^2\right\}$$

Putting the above results together, we have shown that for arbitrary $m \geq 0$,

$$\mathbb{E}\left[\left|\theta - \widehat{\theta}_n\right|^2\right]$$

$$\leq \frac{1}{2\pi^2}\left\{C_1 \int_{\{|u|>\pi m\}} |\mathcal{F}f(-u)|^2\,\mathrm{d}u \wedge C_2\left(\int |\mathcal{F}f(-u)|\,\mathrm{d}u\right)^2\right\}$$

$$+ \frac{T^{-1}}{2\pi^2}\left\{C_1 \int_{\{|u|\leq\pi m\}} \frac{|\mathcal{F}f(-u)|^2}{|\varphi_{\Delta}(u)|^2}\,\mathrm{d}u \wedge C_2\left(\int_{\{|u|\leq\pi m\}} \frac{|\mathcal{F}f(-u)|}{|\varphi_{\Delta}(u)|}\,\mathrm{d}u\right)^2\right\},$$

which is the statement of the theorem. □

5.2 Rate results

There is no need to give a detailed discussion on the rate results since the bounds on the bias terms are standard analysis and the esitimate of the variance term and the minimzation problems leading to Theorm 2.11 and 2.12 are trivial. The proofs are given in full length in [17].

5.3 Proof of Theorem 3.3

5.3.1 Preliminaries

We start by restating for the reader's convenience, some well known results.

5.2 Lemma (Bernstein's Inequality)**.** *Let $X_1, \cdots, X_n$ be complex valued i.i.d. random variables with $\operatorname{Var}(X_1) \leq v^2$ and suppose that $\|X_1\|_\infty \leq B$ for some $B < \infty$. Let $S_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])$. Then the following holds true for arbitrary $\kappa > 0$:*

$$\mathbb{P}\left(\left\{|S_n| \geq \kappa\right\}\right) \leq 4\exp\left(-n\frac{\kappa^2}{4v^2 + \frac{4\sqrt{2}}{3}\kappa B}\right)$$

A proof of this fundamental result can be found, for example in [11].

The following integral version of the classical Berstein inequality can be derived readily from Lemma 5.2:

5.3 Lemma. *In the situation of the preceding lemma, suppose that $\mathbb{E}[|S_n|] \leq H$. Then we have*

$$\mathbb{E}\left[\left\{|S_n|^2 - H^2\right\}_+\right] \leq 32\frac{v^2}{n}\exp\left(-n\frac{H^2}{8v^2}\right) + 128\sqrt{2}\frac{B^2}{n^2}\exp\left(-n\frac{H}{\frac{16\sqrt{2}}{3}B}\right).$$

Finally, we need the Talagrand inequality, which strengthens the classical Bernstein inequality to countable sets of random variables:

5.4 Lemma (Talagrand's inequality)**.** *Let I be some countable index set. For each $i \in I$, let $X_1^{(i)}, \cdots, X_n^{(i)}$ be centered i.i.d. complex valued random variables with $\|X_1^{(i)}\|_\infty \leq B$ for some $B < \infty$. Let $v^2 := \sup_{i\in I}\operatorname{Var} X_1^{(i)}$. Then for arbitrary $\varepsilon > 0$, there are positive constants c_1 and $c_2(\varepsilon)$ depending only on ε such that for any $\kappa > 0$:*

$$\mathbb{P}\left(\left\{\sup_{i\in I}|S_n^{(i)}| \geq (1+\varepsilon)\,\mathbb{E}\left[\sup_{i\in I}|S_n^{(i)}|\right] + \kappa\right) \leq 4\exp\left(-n\frac{\kappa^2}{c_1 v^2 + c_2(\varepsilon)\kappa B}\right\}\right)$$

5.3.2 Auxiliary results

In what follows, we formulate and prove a number of auxiliary results which will be needed in order to prove the main result of Section 3.3.

5.5 Lemma. *Let $\tau > 0$ be given. Let δ be the constant appearing in the definition of the weight function w and let c_1 be the constant in Talagrand's inequality. Then, for arbitrary $\gamma > 0$, there is a positive constant $C_K = C_K^{\tau,\gamma,\delta}$ depending on the choice of τ, γ and δ such that we have for $n \geq 1$:*

$$\mathbb{P}\left(\left\{\exists u \in \mathbb{R} : |\widehat{\varphi}_n(u) - \varphi(u)| \geq \tau(\log n)^{1/2} w(u)^{-1} n^{-1/2}\right\}\right) \leq C_K n^{-\frac{(\tau-\gamma)^2}{c_1}}$$

Proof. We proof the claim for the countable set of rational numbers. By continuity of the characteristic function and of w, it carries over to the whole range of real numbers.

By Theorem 4.1 in [19], we have for some constant C_{RN}:

$$\mathbb{E}\left[\sup_{u\in\mathbb{R}} |\widehat{\varphi}_n(u) - \varphi(u)| w(u)\right] \leq C_{RN} n^{-1/2}.$$

Since moreover, we trivially have $\sup_{u\in\mathbb{R}} \mathrm{Var}[\widehat{\varphi}_1(u)] \leq 1$, and $\sup_{u\in\mathbb{R}} \|\widehat{\varphi}_1(u) w(u)\|_\infty \leq 1$, we can apply Talagrand's inequality. Setting

$$\kappa_n := \tau(\log n)^{1/2} n^{-1/2} - (1+\varepsilon) C_{RN} n^{-1/2},$$

for some $\varepsilon > 0$, we can estimate

$$\begin{aligned}
& \mathbb{P}\left(\left\{\exists q \in \mathbb{Q} : |\widehat{\varphi}_n(q) - \varphi(q)| \geq \tau(\log n)^{1/2} w(q)^{-1} n^{-1/2}\right\}\right) \\
= \;& \mathbb{P}\left(\left\{\sup_{q\in\mathbb{Q}} |\widehat{\varphi}_n(q) - \varphi(q)| w(q) \geq \tau(\log n)^{1/2} n^{-1/2}\right\}\right) \\
\leq \;& \mathbb{P}\left(\left\{\sup_{q\in\mathbb{Q}} |\widehat{\varphi}_n(q) - \varphi(q)| w(q) \geq (1+\varepsilon)\, \mathbb{E}\left[\sup_{q\in\mathbb{Q}} |\widehat{\varphi}_n(q) - \varphi(q)| w(q)\right] + \kappa_n\right\}\right) \\
\leq \;& 4 \exp\left(-n \frac{\kappa_n^2}{c_1 + c_2(\varepsilon)\kappa_n}\right).
\end{aligned}$$

By definition of κ_n, we have for C_K large enough and arbitrary $n \geq 1$:

$$4 \exp\left(-n \frac{\kappa_n^2}{c_1 + c_2(\varepsilon)\kappa_n}\right) \leq C_K \exp\left(-\frac{(\tau-\gamma)^2}{c_1} (\log n)\right) = C_K n^{-\frac{(\tau-\gamma)^2}{c_1}}$$

This is the desired result for the rational numbers and hence, by continuity, for the real line. $\square$

We can now use Lemma 5.5 to analyse the deviation of $\frac{1}{\tilde{\widehat{\varphi}}_n}$ from $\frac{1}{\varphi}$.

5.6 Proposition. *Let* $\frac{1}{\widetilde{\widetilde{\varphi}}_n} = \frac{1}{\widetilde{\widetilde{\varphi}}_n^{\kappa,\delta}}$ *be defined as in Definition (3.1). Assume that for some* $\gamma > 0$ *and some* $p > 0$*, we have* $\kappa \geq 2(\sqrt{pc_1} + \gamma)$*, where* c_1 *denotes the constant in Talagrands inequality. Then we find that*

$$\mathbb{P}\left(\left\{\exists u \in \mathbb{R} : \left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2 > \left((4\kappa)^2 \frac{(\log n) w(u)^{-1} n^{-1}}{|\varphi(u)|^4} \wedge \left(\frac{5}{2}\right)^2 \frac{1}{|\varphi(u)|^2}\right)\right\}\right)$$
$$= \quad O\left(n^{-p}\right).$$

Proof. (Sketch) Let us introduce the favourable set

$$C := C^{\kappa,\delta} := \left\{\forall u \in \mathbb{R} : |\widehat{\varphi}_n(u) - \varphi(u)| \leq \frac{\kappa}{2}(\log n)^{1/2} w(u)^{-1} n^{-1/2}\right\}.$$

We start by recalling that, thanks to Lemma 5.5, we have,

$$\mathbb{P}\left(C^c\right) \leq C_K n^{-\frac{(\kappa/2-\gamma)^2}{c_1}} = O\left(n^{-p}\right),$$

so it is enough to consider the set C.

Let us introduce the following partition of the real line: We have $\mathbb{R} = \mathbb{R}_1^\kappa \cup \mathbb{R}_2^\kappa \cup \mathbb{R}_3^\kappa$, with

$$\mathbb{R}_1^\kappa = \left\{u \in \mathbb{R} : |\varphi(u)| < \frac{\kappa}{2}(\log n)^{1/2} w(u)^{-1} n^{-1/2}\right\},$$
$$\mathbb{R}_2^\kappa = \left\{u \in \mathbb{R} : |\varphi(u)| > \frac{3}{2}\kappa(\log n)^{1/2} w(u)^{-1} n^{-1/2}\right\}$$

and

$$\mathbb{R}_3^\kappa = \left\{u \in \mathbb{R} : \frac{\kappa}{2}(\log n)^{1/2} w(u)^{-1} n^{-1/2} \leq |\varphi(u)| \leq \frac{3}{2}\kappa(\log n)^{1/2} w(u)^{-1} n^{-1/2}\right\}.$$

We prove the claim for each of these sets separately. Since the details are elementary and rely on an repeated application of Lemma 5.5, we only give the details for $\mathbb{R}_1^\kappa$ ond omitt the rest of the proof.

By definition of C, we find that for arbitrary $u \in \mathbb{R}_1^\kappa$, we have

$$|\widehat{\varphi}_n(u)| \leq |\varphi(u)| + |\varphi(u) - \widehat{\varphi}_n(u)| < \kappa(\log n)^{1/2} w(u)^{-1} n^{-1/2} \tag{5.32}$$

and hence by definition of $\frac{1}{\widetilde{\widetilde{\varphi}}_n}$:

$$\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2 = \frac{|\varphi(u) - \widetilde{\widetilde{\varphi}}_n(u)|^2}{|\varphi(u)|^2 |\widetilde{\widetilde{\varphi}}_n(u)|^2} \tag{5.33}$$
$$= \frac{|\varphi(u) - \kappa(\log n)^{1/2} w(u)^{-1} n^{-1/2}|^2}{|\widetilde{\widetilde{\varphi}}_n(u)|^2 |\varphi(u)|^2} \tag{5.34}$$
$$\leq \frac{\left(\frac{3}{2}\kappa\right)^2 (\log n) w(u)^{-2} n^{-1}}{|\varphi(u)|^2 |\widetilde{\widetilde{\varphi}}_n(u)|^2}. \tag{5.35}$$

Next, notice that, using again the definition of C and $\mathbb{R}_1^\kappa$,

$$\frac{1}{|\widetilde{\widetilde{\varphi}}_n(u)|^2} = \kappa^{-2}(\log n)^{-1}w(u)^2 n \tag{5.36}$$

$$\leq \frac{1}{4}\left(\frac{\kappa}{2}\right)^{-2}(\log n)^{-1}w(u)^2 n \leq \frac{1}{4}\frac{1}{|\varphi(u)|^2}. \tag{5.37}$$

Putting (5.33)- (5.37) together, we have shown that on C, we have for any $u \in \mathbb{R}_1^\kappa$:

$$\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2 \leq \left(\frac{3}{4}\kappa\right)^2 \frac{(\log n)w(u)^{-2}n^{-1}}{|\varphi(u)|^4} \wedge \frac{9}{4}\frac{1}{|\varphi(u)|^2}.$$

Similar arguments can be applied to prove the claim for $u \in \mathbb{R}_2^\kappa$ and $u \in \mathbb{R}_3^\kappa$. □

The following result can be derived immediately from the preceding statement.

5.7 Corollary. *In the situation of the preceding statement, we have*

$$\mathbb{P}\left(\left\{\exists u \in \mathbb{R} : \left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2 > \left(\frac{5}{2}\kappa\right)^2 \frac{(\log n)w(u)^{-2}n^{-1}}{|\widetilde{\widetilde{\varphi}}_n(u)|^2|\varphi(u)|^2}\right\}\right) = O\left(n^{-p}\right). \tag{5.38}$$

This corollary is an immediate consequence of the proof of the preceding statement, see lines (5.33)- (5.34).

Lemma 3.2 can now be stated as a consequence of Proposition 5.6:

Proof of Lemma 3.2. Let the set C be defined as in the proof of Proposition 5.6. We can decompose

$$\mathbb{E}\left[\sup_{u\in\mathbb{R}} \frac{\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2}{\frac{(\log n)w(u)^{-2}n^{-1}}{|\varphi(u)|^4} \wedge \frac{1}{|\varphi(u)|^2}}\right] \tag{5.39}$$

$$\leq \mathbb{E}\left[\sup_{u\in\mathbb{R}} \frac{\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2}{\frac{(\log n)w(u)^{-2}n^{-1}}{|\varphi(u)|^4} \wedge \frac{1}{|\varphi(u)|^2}} 1(C)\right] \tag{5.40}$$

$$+ \mathbb{E}\left[\sup_{u\in\mathbb{R}} \frac{\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2}{\frac{(\log n)w(u)^{-2}n^{-1}}{|\varphi(u)|^4} \wedge \frac{1}{|\varphi(u)|^2}} 1(C^c)\right] \tag{5.41}$$

The definition of C, together with Proposition 5.6, readily implies that

$$\mathbb{E}\left[\sup_{u\in\mathbb{R}} \frac{\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2}{\frac{(\log n)w(u)^{-2}n^{-1}}{|\varphi(u)|^4} \wedge \frac{1}{|\varphi(u)|^2}} 1(C)\right] \leq \frac{25}{4}\kappa^2. \tag{5.42}$$

On the other hand, since $\frac{1}{|\widetilde{\widetilde{\varphi}}_n(u)|} \leq \kappa^{-1}(\log n)^{-1/2}w(u)n^{1/2}$ by definition, we can always estimate

$$\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2 \left((\log n)^{-1}w(u)^2 n|\varphi(u)|^4 \vee |\varphi(u)|^2\right) \leq 2n^2, \tag{5.43}$$

which yields

$$\mathbb{E}\left[\sup_{u\in\mathbb{R}} \frac{\left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2}{\frac{(\log n)w(u)^{-2}n^{-1}}{|\varphi(u)|^4} \wedge \frac{1}{|\varphi(u)|^2}} 1(C^c)\right] \leq 2n^2\, \mathbb{P}\left(C^c\right), \tag{5.44}$$

and this expression is bounded by some constant since $\mathbb{P}(C^c) = O(n^{-2})$ by assumption on κ and by Lemma 5.5. □

We conclude this section by formulating two more auxiliary resuls:

5.8 Lemma. *For some $\gamma > 0$, let $\kappa = 2\left(\sqrt{2pc_1} + \gamma\right)$. Let*

$$\begin{aligned} x^2_{m,k,f,w} \quad := \quad & \frac{1}{2\pi^2}\Bigg\{\bar{C}_1 \int |\mathcal{F}f(-u)|^2 \left|\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right|^2 w(u)^{-2}\, du \\ & \wedge \bar{C}_2 \left(\int |\mathcal{F}f(-u)| \left|\mathcal{F}\mathrm{K}\left(\frac{u}{k}\right) - \mathcal{F}\mathrm{K}\left(\frac{u}{m}\right)\right| w(u)^{-1}\, du\right)^2\Bigg\} \end{aligned}$$

and $\lambda_{m,k,f,w} := \log\left(x^2_{m,k,f,w}(k-m)^2\right)$.

Then we have for some constant C_K depending on γ:

$$\begin{aligned} & \mathbb{P}\left(\left\{\exists u \in \mathbb{R} : |\widehat{\varphi}_n(u) - \varphi(u)| \geq \left(\frac{\kappa}{2}(\log n)^{1/2} + \lambda_{m,k,f,w}\right) w(u)^{-1}n^{-1/2}\right\}\right) \\ \leq \quad & C_K n^{-p} n^{-p} x^{-2}_{m,k,f,w}(k-m)^{-2}. \end{aligned}$$

Proof. The proof runs exactly along the same lines as the proof of Lemma 5.5, setting, this time

$$\kappa_n := \left(\frac{\kappa}{2}(\log n)^{1/2} + \lambda_{m,k,f,w}\right) n^{-1/2} - C_{NR}n^{-1/2}.$$

□

5.9 Lemma. *In the situation of the preceding statement, we have*

$$\mathbb{P}\left(\left\{\exists u\in\mathbb{R}:\left|\frac{1}{\varphi(u)}-\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)}\right|^2>\frac{\left(\frac{5}{2}\kappa(\log n)^{1/2}+\lambda_{m,k,f,w}\right)^2 w(u)^{-2}n^{-1}}{|\widetilde{\widetilde{\varphi}}_n(u)|^2|\varphi(u)|^2}\right\}\right)$$
$$\leq\ C_K n^{-p}x_{m,k,f,w}^{-2}(k-m)^{-2}.$$

This statement is derived from Lemma 5.8, using the same arguments which are given to derive Corollary 5.7.

5.3.3 Preparing the proof of the main result

To be able to prove the main result of this section, we will need the following auxiliary result:

5.10 Proposition. *For arbitrary $m\in\mathbb{N}$, we can estimate*

$$\mathbb{E}\left[\sup_{\substack{k>m\\k\in\mathcal{M}}}\left\{\left|\left(\widehat{\theta}_k-\widehat{\theta}_m\right)-(\theta_k-\theta_m)\right|^2-\frac{1}{2}\widetilde{H}^2(m,k)\right\}_+\right]=O\left(n^{-1}\right).$$

Proof. (*Sketch*) The proof of this statement is long, but the steps are elementary. We content ourselves with giving the main ideas. For the details, we refer to [17].

For $m\in\mathcal{M}$, let

$$\widetilde{\theta}_m:=\frac{1}{2\pi}\int\mathcal{F}f(-u)\frac{\varphi'(u)}{\widetilde{\widetilde{\varphi}}_n(u)}\mathcal{F}\,\mathrm{K}\left(\frac{u}{m}\right)\,\mathrm{d}u.$$

We use the estimate

$$\left|\left(\widehat{\theta}_k-\widehat{\theta}_m\right)-(\theta_k-\theta_m)\right|^2$$
$$\leq\ 2\left|\left(\widehat{\theta}_k-\widehat{\theta}_m\right)-\left(\widetilde{\theta}_k-\widetilde{\theta}_m\right)\right|^2+2\left|\left(\widetilde{\theta}_k-\widetilde{\theta}_m\right)-(\theta_k-\theta_m)\right|^2.$$

First, we show that

$$\mathbb{E}\left[\sup_{\substack{k>m\\k\in\mathbb{N}}}\left\{\left|(\widehat{\theta}_k-\widehat{\theta}_m)-(\widetilde{\theta}_k-\widetilde{\theta}_m)\right|^2-\frac{1}{8}\widetilde{\mathrm{H}}^2(m,k)\right\}_+\right]$$
$$\leq\ \sum_{\substack{k\geq m\\k\in\mathcal{M}}}\mathbb{E}\left[\left\{\left|(\widehat{\theta}_k-\widehat{\theta}_m)-(\widetilde{\theta}_k-\widetilde{\theta}_m)\right|^2-\frac{1}{8}\widetilde{\mathrm{H}}^2(m,k)\right\}_+\right]$$

is negligible. This is done by applying the integral version of Bernstein's inequality to the conditional expectations

$$\mathbb{E}\left[\left\{\left|(\widehat{\theta}_k-\widehat{\theta}_m)-(\widetilde{\theta}_k-\widetilde{\theta}_m)\right|^2-\frac{1}{8}\widetilde{\mathrm{H}}^2(m,k)\right\}_+\middle|\widehat{\varphi}_n\right]$$

and concluding that the sum is (almost surely) negligible.

Since $\left(\widehat{\theta}_k - \widehat{\theta}_m\right)$ is clearly unbounded, one has to truncate the random variables $Z_j = X_j - X_{j-1}$ at the threshold $\log n + \log\left(x_{m,k}^2 (m-k)^2\right)$. Then one can directly apply Lemma 5.3 to see that the sum is negligible.

The remainder terms are seen to be negligible, using the Markov inequality. To do this, we need the exponential moment condition on X_1.

Next, we consider

$$\mathbb{E}\left[\sup_{\substack{k>m\\ k\in\mathbb{N}}}\left\{\left|(\widetilde{\theta}_k - \widetilde{\theta}_m) - (\theta_k - \theta_m)\right|^2 - \frac{1}{8}\widetilde{\mathrm{H}}^2(m,k)\right\}_+\right].$$

To see that this term is negligible, we first introduce the favourable sets

$$C(m,k) := \left\{\forall u \in \mathbb{R} : \left|\frac{1}{\widetilde{\widetilde{\varphi}}_n(u)} - \frac{1}{\varphi(u)}\right|^2 \leq \frac{\left(\left(\frac{5}{2}\kappa\right)(\log n) + \lambda_{m,k,f,w}\right)^2 w(u)^{-1} n^{-1}}{\left|\widetilde{\widetilde{\varphi}}_n(u)\right|^2 |\varphi(u)|^2}\right\}$$

and show that on $C(m,k)$:

$$\left|(\widehat{\theta}_k - \widehat{\theta}_m) - (\theta_k - \theta_m)\right|^2 \leq \frac{1}{8}\widetilde{\mathrm{H}}^2(m,k).$$

This inequality can be derived immediately from the definition of $C(m,k)$ by a repeated application of the Cauchy-Schwarz inequality.

Finally, it remains to show that

$$\sum_{\substack{k>m\\ k\in\mathcal{M}}}\mathbb{E}\left[\left\{\left|(\widetilde{\theta}_k - \widetilde{\theta}_m) - (\theta_k - \theta_m)\right|^2 - \frac{1}{8}\widetilde{\mathrm{H}}^2(m,k)\right\}_+ \mathbb{1}\left(C(m,k)^c\right)\right]$$

is negligible. This fact can be derived from Lemma 5.9. □

5.3.4 Proof of the main result

We are now ready to prove the oracle inequality given in Theorem 3.3 and thus the key result of our adapitve estimation procedure:

Proof of Theorem 3.3. In what follows, let m^* be the oracle cutoff, that is,

$$m^* = \operatorname*{arginf}_{m\in\mathcal{M}}\left\{\sup_{\substack{k>m\\ k\in\mathcal{M}}}|\theta_k - \theta_m|^2 + \mathrm{pen}(m)\right\}.$$

We start by considering the loss on the set $\{\widehat{m} \leq m^*\}$.

We can decompose

$$\left|\theta - \widehat{\theta}_{\widehat{m}}\right|^2 \leq 2\left|\theta - \widehat{\theta}_{m^*}\right|^2 + 2\left|\widehat{\theta}_{m^*} - \widehat{\theta}_{\widehat{m}}\right|^2.$$

By definition of $\widehat{m}$, we can estimate

$$\begin{aligned} & \left|\widehat{\theta}_{m^*} - \widehat{\theta}_{\widehat{m}}\right|^2 1\left(\{\widehat{m} \leq m^*\}\right) \\ \leq\ & \sup_{\substack{k \geq m^* \\ k \in \mathcal{M}}} \left\{\left|\widehat{\theta}_k - \widehat{\theta}_{m^*}\right|^2 - \widetilde{\mathrm{H}}^2(m^*, k)\right\} + \widetilde{\mathrm{pen}}(m^*) + \widetilde{\mathrm{H}}^2(\widehat{m}, m^*) 1\left(\{\widehat{m} \leq m^*\right). \end{aligned}$$

The definition of $\widetilde{\mathrm{H}}(\widehat{m}, m^*)$ implies that we have

$$\widetilde{\mathrm{H}}^2(\widehat{m}, m^*) 1\left(\{\widehat{m} \leq m^*\}\right) \leq \widetilde{\mathrm{pen}}(m^*) \tag{5.45}$$

and an application of Lemma 3.2 readily implies that

$$\mathbb{E}\left[\widetilde{\mathrm{pen}}(m^*)\right] \leq \mathrm{pen}(m^*) O(1). \tag{5.46}$$

Finally, we can estimate

$$\begin{aligned} & \sup_{\substack{k \geq m^* \\ k \in \mathcal{M}}} \left\{\left|\widehat{\theta}_k - \widehat{\theta}_{m^*}\right|^2 - \widetilde{\mathrm{H}}^2(m^*, k)\right\} \qquad (5.47) \\ \leq\ & 2\sup_{\substack{k \geq m^* \\ k \in \mathcal{M}}} \left\{\left|(\widehat{\theta}_k - \widehat{\theta}_{m^*}) - (\theta_k - \theta_{m^*})\right|^2 - \frac{1}{2}\widetilde{\mathrm{H}}^2(m^*, k)\right\} + 2\sup_{\substack{k \geq m^* \\ k \in \mathcal{M}}} |\theta_k - \theta_{m^*}|^2. \end{aligned}$$

Taking expectation, we find that the first expression appearing in the last line of (5.47) is negligible thanks to Proposition 5.10. Using this, (5.45) and (5.46), we have shown that for some consant C,

$$\begin{aligned} & \mathbb{E}\left[|\theta - \widehat{\theta}_{\widehat{m}}|^2 1\left(\{\widehat{m} \geq m^*\}\right)\right] \\ \leq\ & C \inf_{m \in \mathcal{M}} \left\{|\theta - \theta_{m_n}|^2 + \sup_{\substack{k \geq m \\ k \in \mathcal{M}}} |\theta_k - \theta_m|^2 + \mathrm{pen}(m)\right\} + O\left(n^{-1}\right). \end{aligned}$$

It remains to consider the loss on $\{\widehat{m} > m^*\}$.

We use the decomposition

$$|\theta - \widehat{\theta}_{\widehat{m}}|^2 \leq 2|\theta - \theta_{\widehat{m}}|^2 + 2|\theta_{\widehat{m}} - \widehat{\theta}_{\widehat{m}}|^2.$$

First, we can immediately estimate

$$|\theta - \theta_{\widehat{m}}|^2 1\left(\{\widehat{m} > m^*\}\right) \leq 3\left(|\theta - \theta_{m_n}|^2 + \sup_{\substack{k > m^* \\ k \in \mathcal{M}}} |\theta_k - \theta_{m^*}|^2\right).$$

Next, we can decompose

$$
\begin{aligned}
& |\theta_{\widehat{m}} - \widehat{\theta}_{\widehat{m}}|^2 1\left(\{\widehat{m} > m^*\}\right) && (5.48)\\
\leq & \sum_{\substack{k>m^* \\ k\in\mathcal{M}}} \left\{|\theta_k - \widehat{\theta}_k|^2 - \widetilde{\text{pen}}(k)\right\}_+ + \sum_{\substack{k>m^* \\ k\in\mathcal{M}}} \widetilde{\text{pen}}(k) 1\left(\{\widehat{m} = k\}\right).
\end{aligned}
$$

Again, using Proposition 5.10, we find that the expected value of the first expression appearing in the second line of (5.48) is readily negligible.

Next, we use the fact that by definition of $\widehat{m}$, we have on $\{\widehat{m} = k\}$:

$$
\begin{aligned}
\widetilde{\text{pen}}(k) &\leq \sup_{\substack{l>m^* \\ l\in\mathcal{M}}} \left\{|\widehat{\theta}_l - \widehat{\theta}_{m^*}|^2 - \widetilde{\mathrm{H}}^2(m^*, l)\right\} + \widetilde{\text{pen}}(m^*) \\
&\leq 2 \sup_{\substack{l>m^* \\ l\in\mathcal{M}}} \left\{|(\widehat{\theta}_l - \widehat{\theta}_{m^*}) - (\theta_l - \theta_{m^*})|^2 - \frac{1}{2}\widetilde{\mathrm{H}}^2(m^*, l)\right\} \\
&+ 2 \sup_{\substack{l>m^* \\ l\in\mathcal{M}}} |\theta_l - \theta_{m^*}|^2 + \widetilde{\text{pen}}(m^*)
\end{aligned}
$$

to see that

$$
\begin{aligned}
& \sum_{\substack{k>m^* \\ k\in\mathcal{M}}} \widetilde{\text{pen}}(k) 1\left(\{\widehat{m} = k\}\right) && (5.49)\\
\leq & \; 2 \sup_{\substack{l>m^* \\ l\in\mathcal{M}}} \left\{|(\widehat{\theta}_l - \widehat{\theta}_{m^*}) - (\theta_l - \theta_{m^*})|^2 - \frac{1}{2}\widetilde{\mathrm{H}}^2(m^*, l)\right\} && (5.50)\\
+ & \; 2 \sup_{\substack{l>m^* \\ l\in\mathcal{M}}} |\theta_l - \theta_{m^*}|^2 + \widetilde{\text{pen}}(m^*). && (5.51)
\end{aligned}
$$

Again, we see that the second line in (5.49) is negligible and we use, once more, Lemma 3.2 to see that

$$
\mathbb{E}\left[\widetilde{\text{pen}}(m^*)\right] \leq O(1)\, \text{pen}(m^*).
$$

We have thus shown that for some constant C,

$$
\begin{aligned}
& \mathbb{E}\left[|\theta - \widehat{\theta}_{\widehat{m}}|^2 1\left(\{\widehat{m} > m^*\}\right)\right] \\
\leq & \; C \inf_{m\in\mathcal{M}} \left\{ |\theta - \theta_{m_n}|^2 + \sup_{\substack{k>m \\ k\in\mathcal{M}}} |\theta_k - \theta_m|^2 + \text{pen}(m) \right\} + O(n^{-1}).
\end{aligned}
$$

and hence the main result of Section 3.3). $\square$

References

[1] Denis Belomestny and Markus Reiß. Spectral calibration of exponential Lévy models. *Finance Stoch.*, 10(4):449–474, 2006.

[2] Lucien Birgé. An alternative point of view on Lepski's method. In *State of the art in probability and statistics: Festschrift for Willem R. van Zwet*, pages 113–133. IMS Lecture Notes-Monograph Series, 1999.

[3] Christina Butucea and Fabienne Comte. Adaptive estimation of linear functionals in the convolution model and applications. *Bernoulli*, 15(1):69–98, 2009.

[4] Peter Carr, Hélyette Geman, Dilip Madan, and Marc Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

[5] Peter Carr, Dilip Madan, and Eric Chang. The variance Gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[6] Fabienne Comte and Valentine Genon-Catalot. Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Processes and their Applications*, 119:4088–4123, 2009.

[7] Fabienne Comte and Valentine Genon-Catalot. Nonparametric adaptive estimation for pure jump Lévy processes. *Annales de l'Institut Henri Poincaré*, 3(46):595–617, 2010.

[8] Fabienne Comte and Valentine Genon-Catalot. Estimation for Lévy processes from high frequency data within a long time interval. *Annals of Statistics*, 39(2):803–837, 2011.

[9] Fabienne Comte and Claire Lacour. Data-driven density estimation in the presence of additive noise with unknown distribution. *Journal of the Royal Statistical Society: Series B*, 73:601–627, 2011.

[10] Rama Cont and Peter Tankov. *Financial modelling with jump processes.* Chapman & Hall/CRC Financial Mathematics Series. Boca Raton, 2004.

[11] Richard M. Dudley. *Uniform central limit theorems.* Cambridge Studies in Advanced Mathematics 63. Cambridge: Cambridge University Press. xiv, 436 p , 2008.

[12] José E. Figueroa-López. Small-time moment asymptotics for Lévy processes. *Stat. Probab. Lett.*, 78(18):3355–3365, 2008.

[13] José E. Figueroa-López. Nonparametric estimation for Lévy models based on discrete sampling. *IMS Lecture Notes of the 3rd E.L. Lehmann Symposium*, 57:117–146, 2009.

[14] José E. Figueroa-López and Christian Houdré. Risk bounds for the non-parametric estimation of Lévy processes. Giné, Evarist (ed.) et al., High dimensional probability. Institute of Mathematical Statistics Lecture Notes - Monograph Series 51, 96-116, 2006.

[15] Shota Gugushvili. Nonparametric estimation of the characteristic triplet of a discretely observed Lévy process. *Journal of Nonparametric Statistics*, 21(3):321–343, 2009.

[16] Shota Gugushvili. Nonparametric estimation for discretely sampled Lévy processes. *Annales de l'Institut Henri Poincaré*, 48(1):282–307, 2012.

[17] J. Johanna Kappus. *Nonparametric adaptive estimation for discretely observed Lévy processes*. PhD thesis, Humboldt Universität zu Berlin, 2012. (to appear).

[18] Béatrice Laurent, Carenne Ludeña and Clémentine Prieur. Adaptive estimation of linear functionals by model selection. *Electronic Journal of Statistics*, 2:993–1020, 2008.

[19] Michael Neumann and Markus Reiß. Nonparametric estimation for Lévy processes from low frequency observations. *Bernoulli*, 15(1):223–248, 2009.

[20] Michael H. Neumann. On the effect of estimating the error density in nonparametric deconvolution. *Journal of Nonparametric Statistics*, 7(4):307–330, 1997.

[21] Peter Carr and Hélyette Geman and Dilip Madan and Marc Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13:345–382, 2003.

[22] Peter Carr and Liuren Wu. Time changed Lévy processes and option pricing. *Journal of Financial Economics*, 71:113–141, 2004.

[23] Ken-Iti Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press, 1999.

[24] Jakob Söhl. Confidence sets in nonparametric calibration of exponential Lévy models. SFB 649 Discussion Paper 2012-012, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany, 2012. available at http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2012-012.pdf.

SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "HMM in dynamic HAC models" by Wolfgang Karl Härdle, Ostap Okhrin and Weining Wang, January 2012.
002 "Dynamic Activity Analysis Model Based Win-Win Development Forecasting Under the Environmental Regulation in China" by Shiyi Chen and Wolfgang Karl Härdle, January 2012.
003 "A Donsker Theorem for Lévy Measures" by Richard Nickl and Markus Reiß, January 2012.
004 "Computational Statistics (Journal)" by Wolfgang Karl Härdle, Yuichi Mori and Jürgen Symanzik, January 2012.
005 "Implementing quotas in university admissions: An experimental analysis" by Sebastian Braun, Nadja Dwenger, Dorothea Kübler and Alexander Westkamp, January 2012.
006 "Quantile Regression in Risk Calibration" by Shih-Kang Chao, Wolfgang Karl Härdle and Weining Wang, January 2012.
007 "Total Work and Gender: Facts and Possible Explanations" by Michael Burda, Daniel S. Hamermesh and Philippe Weil, February 2012.
008 "Does Basel II Pillar 3 Risk Exposure Data help to Identify Risky Banks?" by Ralf Sabiwalsky, February 2012.
009 "Comparability Effects of Mandatory IFRS Adoption" by Stefano Cascino and Joachim Gassen, February 2012.
010 "Fair Value Reclassifications of Financial Assets during the Financial Crisis" by Jannis Bischof, Ulf Brüggemann and Holger Daske, February 2012.
011 "Intended and unintended consequences of mandatory IFRS adoption: A review of extant evidence and suggestions for future research" by Ulf Brüggemann, Jörg-Markus Hitz and Thorsten Sellhorn, February 2012.
012 "Confidence sets in nonparametric calibration of exponential Lévy models" by Jakob Söhl, February 2012.
013 "The Polarization of Employment in German Local Labor Markets" by Charlotte Senftleben and Hanna Wielandt, February 2012.
014 "On the Dark Side of the Market: Identifying and Analyzing Hidden Order Placements" by Nikolaus Hautsch and Ruihong Huang, February 2012.
015 "Existence and Uniqueness of Perturbation Solutions to DSGE Models" by Hong Lan and Alexander Meyer-Gohde, February 2012.
016 "Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes" by Johanna Kappus, February 2012.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".